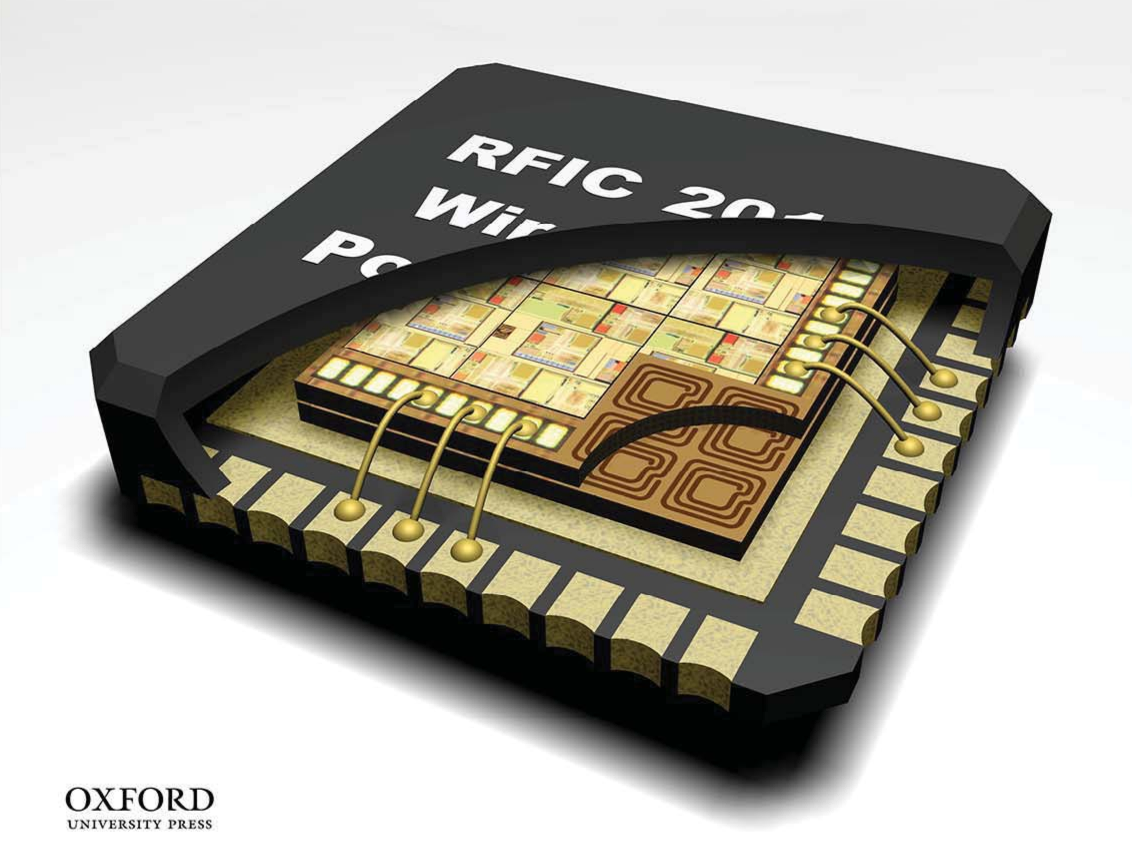
SEDRA/SMITH

Microelectronic Circuits

SEVENTH EDITION



OXFORD
UNIVERSITY PRESS

Microelectronic Circuits

THE OXFORD SERIES IN ELECTRICAL AND COMPUTER ENGINEERING

Adel S. Sedra, Series Editor

Allen and Holberg, *CMOS Analog Circuit Design, 3rd edition*
Bobrow, *Elementary Linear Circuit Analysis, 2nd edition*
Bobrow, *Fundamentals of Electrical Engineering, 2nd edition*
Campbell, *Fabrication Engineering at the Micro- and Nanoscale, 4th edition*
Chen, *Digital Signal Processing*
Chen, *Linear System Theory and Design, 4th edition*
Chen, *Signals and Systems, 3rd edition*
Comer, *Digital Logic and State Machine Design, 3rd edition*
Comer, *Microprocessor-Based System Design*
Cooper and McGillem, *Probabilistic Methods of Signal and System Analysis, 3rd edition*
Dimitrijev, *Principles of Semiconductor Device, 2nd edition*
Dimitrijev, *Understanding Semiconductor Devices*
Fortney, *Principles of Electronics: Analog & Digital*
Franco, *Electric Circuits Fundamentals*
Ghausi, *Electronic Devices and Circuits: Discrete and Integrated*
Guru and Hiziroğlu, *Electric Machinery and Transformers, 3rd edition*
Houts, *Signal Analysis in Linear Systems*
Jones, *Introduction to Optical Fiber Communication Systems*
Krein, *Elements of Power Electronics*
Kuo, *Digital Control Systems, 2nd edition*
Lathi, *Linear Systems and Signals, 2nd edition*
Lathi and Ding, *Modern Digital and Analog Communication Systems, 4th edition*
Lathi, *Signal Processing and Linear Systems*
Martin, *Digital Integrated Circuit Design*
Miner, *Lines and Electromagnetic Fields for Engineers*
Parhami, *Computer Architecture*
Parhami, *Computer Arithmetic, 2nd edition*
Roberts and Sedra, *SPICE, 2nd edition*
Roberts, Taenzler, and Burns, *An Introduction to Mixed-Signal IC Test and Measurement, 2nd edition*
Roulston, *An Introduction to the Physics of Semiconductor Devices*
Sadiku, *Elements of Electromagnetics, 6th edition*
Santina, Stubberud, and Hostetter, *Digital Control System Design, 2nd edition*
Sarma, *Introduction to Electrical Engineering*
Schaumann, Xiao, and Van Valkenburg, *Design of Analog Filters, 3rd edition*
Schwarz and Oldham, *Electrical Engineering: An Introduction, 2nd edition*
Sedra and Smith, *Microelectronic Circuits, 7th edition*
Stefani, Shahian, Savant, and Hostetter, *Design of Feedback Control Systems, 4th edition*
Tsividis/McAndrew, *Operation and Modeling of the MOS Transistor, 3rd edition*
Van Valkenburg, *Analog Filter Design*
Warner and Grung, *Semiconductor Device Electronics*
Wolovich, *Automatic Control Systems*
Yariv and Yeh, *Photonics: Optical Electronics in Modern Communications, 6th edition*
Żak, *Systems and Control*

SEVENTH EDITION

Microelectronic Circuits

Adel S. Sedra
University of Waterloo

Kenneth C. Smith
University of Toronto

New York Oxford
OXFORD UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

For titles covered by Section 112 of the US Higher Education Opportunity Act, please visit www.oup.com/us/he for the latest information about pricing and alternate formats.

Published in the United States of America by
Oxford University Press
198 Madison Avenue, New York, NY 10016
http://www.oup.com

Library of Congress Cataloging-in-Publication Data

Sedra, Adel S., author.
Microelectronic circuits / Adel S. Sedra, University of Waterloo, Kenneth C. Smith, University of Toronto. — Seventh edition.
pages cm. — (The Oxford series in electrical and computer engineering)
Includes bibliographical references and index.
ISBN 978–0–19–933913–6
1. Electronic circuits. 2. Integrated circuits. I. Smith, Kenneth C. (Kenneth Carless), author. II. Title.
TK7867.S39 2014
621.3815—dc23 2014033965

Cover Photo: This 3D IC system demonstrates the concept of wireless power delivery and communication through multiple layers of CMOS chips. The communication circuits were demonstrated in an IBM 45 nm SOI CMOS process. This technology is designed to serve a multi-Gb/s interconnect between cores spread across several IC layers for high-performance processors.
(Photo Credit: The picture is courtesy of Professor David Wentzloff, Director of the Wireless Integrated Circuits Group at the University of Michigan, and was edited by Muhammad Faisal, Founder of Movellus Circuits Incorporated.)

Printing number: 9 8 7 6 5 4 3 2 1

Printed in the United States of America
on acid-free paper

BRIEF TABLE OF CONTENTS

CONTENTS

Appendices

TABLES

FOR REFERENCE AND STUDY

"EXPAND-YOUR-PERSPECTIVE" NOTES

PREFACE

Microelectronic Circuits, Seventh Edition, is intended as a text for the core courses in electronic circuits taught to majors in electrical and computer engineering. It should also prove useful to engineers and other professionals wishing to update their knowledge through self-study.

As was the case with the first six editions, the objective of this book is to develop in the reader the ability to analyze and design electronic circuits, both analog and digital, discrete and integrated. While the application of integrated circuits is covered, emphasis is placed on transistor circuit design. This is done because of our belief that even if the majority of those studying this book were not to pursue a career in IC design, knowledge of what is inside the IC package would enable intelligent and innovative application of such chips. Furthermore, with the advances in VLSI technology and design methodology, IC design itself has become accessible to an increasing number of engineers.

Prerequisites

The prerequisite for studying the material in this book is a first course in circuit analysis. As a review, some linear circuits material is included here in the appendices: specifically, two-port network parameters in Appendix C; some useful network theorems in Appendix D; single-time-constant circuits in Appendix E; and s-domain analysis in Appendix F. In addition, a number of relevant circuit analysis problems are included at the beginning of the end-of-chapter problems section of Chapter 1. No prior knowledge of physical electronics is assumed. All required semiconductor device physics is included, and Appendix A provides a brief description of IC fabrication. All these appendices can be found on the book's website.

Emphasis on Design

It has been our philosophy that circuit design is best taught by pointing out the various tradeoffs available in selecting a circuit configuration and in selecting component values for a given configuration. The emphasis on design has been retained in this edition. In addition to design examples, and design-oriented exercises and end-of-chapter problems (indicated with a D), the book includes on its website an extensive appendix (Appendix B) where a large number of simulation and design examples are presented. These emphasize the use of SPICE, the most valuable circuit-design aid.

New to the Seventh Edition

While maintaining the philosophy and pedagogical approach of the first six editions, several changes have been made to both organization and coverage. Our goal in making structural changes has been to increase modularity and thus flexibility for the instructor, without causing disturbance to courses currently using the sixth edition. Changes in coverage are necessitated by the continuing advances in technology which make some topics of greater relevance and others of less interest. As well, advances in IC process technology require that the numbers used in the examples, exercises and end-of-chapter problems be updated to reflect the parameters of newer generations of IC technologies (e.g., some problems utilize the parameters of the 65-nm CMOS process). This ensures that students are acquiring a real-world perspective on technology.

To improve presentation, a number of chapters and sections have been rewritten for greater clarity. Specific, noteworthy changes are:

1. **New End-of-Chapter Problems and a New Instructor's Solutions Manual.** The number of the end-of-chapter problems has increased by about 50. Of the resulting 1532 problems, 176 are entirely new and 790 have new data. The new Instructor's Solutions Manual is written by Adel Sedra.
2. **Expand-Your-Perspective Notes**. This is a new feature providing historical and application perspectives. About two such notes are included in each chapter. Most are focused on notable circuit engineers and key inventions.
3. **Greater Flexibility in Presenting the MOSFET and the BJT.** Two short and completely parallel chapters present the MOSFET (Chapter 5) and the BJT (Chapter 6). Here the focus is on the device structure and its physical operation, its current-voltage characteristics, and its application in dc circuits. The order of coverage of these two chapters is entirely at the instructor's discretion as they have been written to be completely independent of each other.
4. **A Unified Treatment of Transistor Amplifiers.** The heart of a first course in electronics is the study of transistor amplifiers. The seventh edition provides a new approach to this subject: A new Chapter 7 begins with the basic principles that underlie the operation of a transistor of either type as an amplifier, and presents such concepts as small-signal operation and modeling. This is followed by the classical configurations of transistor amplifiers, biasing methods, and practical discrete-circuit amplifiers. The combined presentation emphasizes the unity of the basic principles while allowing for separate treatment of the two device types where this is warranted. Very importantly, we are able to compare the two devices and to draw conclusions about their unique areas of application.
5. **Improved Presentation of Cascoding.** Chapter 8 dealing with the basic building blocks of IC amplifiers has been rewritten to improve presentation. Specifically, the development of cascoding and the key circuit building blocks, the cascode amplifier and the cascode current source, is now much clearer.
6. **Clearer and Simplified Study of Feedback.** The feedback chapter has been rewritten to improve, simplify and clarify the presentation of this key subject.
7. **Streamlined Presentation of Frequency Response.** While keeping the treatment of frequency response all together, the chapter has been rewritten to streamline its flow, and simplify and clarify the presentation.
8. **Updated Treatment of Output Stages and Power Amplifiers.** Here, we have updated the material on MOS power transistors and added a new section on the increasingly important class-D switching power amplifier.
9. **A More Contemporary Approach to Operational Amplifier Circuits.** While maintaining coverage of some of the enduring features and subcircuits of the classical 741 op amp, its total coverage is somewhat reduced to make room for modern IC op amp design techniques.

10. **Better Organized and Modernized Coverage of Digital IC Design**. Significant improvements have been made to the brief but comprehensive coverage of digital IC design in Part III. These include a better motivated study of CMOS logic circuits (Chapter 14) which now begins with logic gate circuits. The material on logic circuit technologies and design methodologies as well as the advanced topic of technology scaling and its implications have been moved to Chapter 15. This modularly structured chapter now deals with a selection of advanced and somewhat specialized topics. Since bipolar is hardly ever used in new digital design, coverage of ECL has been significantly reduced. Similarly, BiCMOS has become somewhat of a specialty topic and its coverage has been correspondingly reduced. Nevertheless, the complete material on both ECL and BiCMOS is now available on the book's website. Finally, we have added a new section on image sensors to Chapter 16 (Memory Circuits).
11. **Increased Emphasis on Integrated-Circuit Filters and Oscillators.** A section on a popular approach to integrated-circuit filter design, namely, Transconductance-C filters, has been added to Chapter 17. To make room for this new material, the subsection on stagger-tuned amplifiers has been removed and placed in Appendix H, on the website. The cross-coupled LC oscillator, popular in IC design, has been added to Chapter 18. The section on precision diode circuits has been removed but is still made available on the website.
12. **A Useful and Insightful Comparison of the MOSFET and the BJT.** This is now included in Appendix G, available on the website.

The Book's Website

A Companion Website for the book has been set up at www.oup.com/us/sedrasmith. Its content will change frequently to reflect new developments. The following material is available on the website:

1. Data sheets for hundreds of useful devices to help in laboratory experiments as well as in design projects.
2. Links to industrial and academic websites of interest.
3. A message center to communicate with the authors and with Oxford University Press.
4. Links to the student versions of both Cadence PSpice® and National Instruments Multisim™.
5. The input files for all the PSpice® and Multisim™ examples of Appendix B.
6. Step-by-step guidance to help with the simulation examples and the end-of-chapter problems identified with a SIM icon.
7. Bonus text material of specialized topics which are either not covered or covered briefly in the current edition of the textbook. These include:
   - Junction Field-Effect Transistors (JFETs)
   - Gallium Arsenide (GaAs) Devices and Circuits
   - Transistor-Transistor Logic (TTL) Circuits
   - Emitter-Coupled Logic (ECL) Circuits
   - BiCMOS Circuits
   - Precision Rectifier Circuits
8. Appendices for the Book:
   - Appendix A: VLSI Fabrication Technology
   - Appendix B: SPICE Device Models and Design and Simulation Examples Using PSpice® and Multisim™
   - Appendix C: Two-Port Network Parameters
   - Appendix D: Some Useful Network Theorems
   - Appendix E: Single-Time-Constant Circuits
   - Appendix F: *s*-domain Analysis: Poles, Zeros, and Bode Plots
   - Appendix G: Comparison of the MOSFET and the BJT

- Appendix H: Design of Stagger-Tuned Amplifiers
- Appendix I: Bibliography
- Appendix L: Answers to Selected Problems

Exercises and End-of-Chapter Problems

Over 475 Exercises are integrated throughout the text. The answer to each exercise is given below the exercise so students can check their understanding of the material as they read. Solving these exercises should enable the reader to gauge his or her grasp of the preceding material. In addition, more than 1530 end-of-chapter Problems, 65% of which are new or revised in this edition, are provided. The problems are keyed to the individual chapter sections and their degree of difficulty is indicated by a rating system: difficult problems are marked with an asterisk (*); more difficult problems with two asterisks (**); and very difficult (and/or time consuming) problems with three asterisks (***). We must admit, however, that this classification is by no means exact. Our rating no doubt depended to some degree on our thinking (and mood!) at the time a particular problem was created. Answers to sample problems are given in Appendix L (on the website), so students have a checkpoint to tell if they are working out the problems correctly. Complete solutions for all exercises and problems are included in the *Instructor's Solutions Manual*, which is available from the publisher to those instructors who adopt the book.

As in the previous six editions, many examples are included. The examples, and indeed most of the problems and exercises, are based on real circuits and anticipate the applications encountered in designing real-life circuits. This edition continues the use of numbered solution steps in the figures for many examples, as an attempt to recreate the dynamics of the classroom.

Course Organization

The book contains sufficient material for a sequence of two single-semester courses, each of 40-50 lecture hours. The modular organization of the book provides considerable flexibility for course design. In the following, we suggest content for a sequence of two classical or standard courses. We also describe some variations on the content of these two courses and specify supplemental material for a possible third course.

The First Course

The first course is based on Part I of the book, that is, Chapters 1–7. It can be taught, most simply by starting at the beginning of Chapter 1 and concluding with the end of Chapter 7. However, as guidance to instructors who wish to follow a different order of presentation or a somewhat modified coverage, or to deal with situations where time might be constrained, we offer the following remarks:

The core of the first course is the study of the two transistor types, Chapters 5 and 6, in whatever order the instructor wishes, and transistor amplifiers in Chapter 7. These three chapters must be covered in full.

Another important part of the first course is the study of diodes (Chapter 4). Here, however, if time does not permit, some of the applications in the later part of the chapter can be skipped.

We have found it highly motivational to cover op amps (Chapter 2) near the beginning of the course. This provides the students with the opportunity to work with a practical integrated circuit and to experiment with non-trivial circuits.

Coverage of Chapter 1, at least of the amplifier sections, should prove helpful. Here the sections on signals can be either covered in class or assigned as reading material. Section 1.6 on frequency response is needed if the frequency-response of op-amp circuits is to be studied; otherwise this section can be delayed to the second course.

Finally, if the students have not taken a course on physical electronics, Chapter 3 needs to be covered. Otherwise, it can be used as review material or skipped altogether.

The Second Course

The main subject of the second course is integrated-circuit amplifiers and is based on Part II of the book, that is, Chapters 8-13. Here also, the course can be taught most simply by beginning with Chapter 8 and concluding with Chapter 13. However, this being a second course, considerable flexibility in coverage is possible to satisfy particular curriculum designs and/or to deal with time constraints.

First, however, we note that the core material is presented in Chapters 8-11 and these four chapters must be covered, though not necessarily in their entirety. For instance, some of the sections near the end of a chapter and identified by the "advanced material" icon can be skipped, usually with no loss of continuity.

Beyond the required chapters, (8-11), the instructor has many possibilities for the remainder of the course. These include one or both of the two remaining chapters in Part II, namely, Output Stages and Power Amplifier (Chapter 12), and Op-Amp Circuits (Chapter 13).

Another possibility, is to include an introduction to digital integrated circuits by covering Chapter 14, and if time permits, selected topics of Chapters 15 and 16.

Yet another possibility for the remainder of the second course is selected topics from the filters chapter (17) and/or the oscillators chapter (18).

A Digitally Oriented First Course

A digitally-oriented first course can include the following: Chapter 1 (without Section 1.6), Chapter 2, Chapter 3 (if the students have not had any exposure to physical electronics), Chapter 4 (perhaps without some of the later applications sections), Chapter 5, selected topics from Chapter 7 emphasizing the basics of the application of the MOSFET as an amplifier, Chapter 14, and selected topics from Chapters 15 and 16. Such a course would be particularly suited for Computer Engineering students.

Supplemental Material/Third Course

Depending on the selection of topics for the first and second courses, some material will remain and can be used for part of a third course or as supplemental material to support student design projects. These can include Chapter 12 (Output Stages and Power Amplifiers), Chapter 13 (Op-Amp Circuits), Chapter 17 (Filters) and Chapter 18 (Oscillators), which can be used to support a third course on analog circuits. These can also include Chapters 14, 15 and 16 which can be used for a portion of a senior-level course on digital IC design.

The Accompanying Laboratory

Courses in electronic circuits are usually accompanied by laboratory experiments. To support the laboratory component for courses using this book, Professor Vincent Gaudet of the University of Waterloo has, in collaboration with K.C. Smith, authored a laboratory manual. *Laboratory Explorations*, together with an Instructor's Manual, is available from Oxford University Press.

Another innovative laboratory instruction system, designed to accompany this book, has been recently developed. Specifically, Illuster Technologies Inc. has developed a digitally controlled lab platform, AELabs. The platform is realized on printed circuit boards using surface mount devices. A wide variety of circuits can be configured on this platform through a custom graphical user interface. This allows students to conduct many experiments relatively quickly. More information is available from Illuster (see link on the Companion Website).

An Outline for the Reader

Part I, *Devices and Basic Circuits*, includes the most fundamental and essential topics for the study of electronic circuits. At the same time, it constitutes a complete package for a first course on the subject.

Chapter 1. The book starts with an introduction to the basic concepts of electronics in Chapter 1. Signals, their frequency spectra, and their analog and digital forms are presented. Amplifiers are introduced as circuit building blocks and their various types and models are studied. This chapter also establishes some of the terminology and conventions used throughout the text.

Chapter 2. Chapter 2 deals with operational amplifiers, their terminal characteristics, simple applications, and practical limitations. We chose to discuss the op amp as a circuit building block at this early stage simply because it is easy to deal with and because the student can experiment with op-amp circuits that perform nontrivial tasks with relative ease and with a sense of accomplishment. We have found this approach to be highly motivating to the student. We should point out, however, that part or all of this chapter can be skipped and studied at a later stage (for instance, in conjunction with Chapter 9, Chapter 11, and/or Chapter 13) with no loss of continuity.

Chapter 3. Chapter 3 provides an overview of semiconductor concepts at a level sufficient for understanding the operation of diodes and transistors in later chapters. Coverage of this material is useful in particular for students who have had no prior exposure to device physics. Even those with such a background would find a review of Chapter 3 beneficial as a refresher. The instructor can choose to cover this material in class or assign it for outside reading.

Chapter 4. The first electronic device, the diode, is studied in Chapter 4. The diode terminal characteristics, the circuit models that are used to represent it, and its circuit applications are presented. Depending on the time available in the course, some of the diode applications (e.g. Section 4.6) can be skipped. Also, the brief description of special diode types (Section 4.7) can be left for the student to read.

Chapters 5 and 6. The foundation of electronic circuits is established by the study of the two transistor types in use today: the MOS transistor in Chapter 5 and the bipolar transistor in Chapter 6. ***These two chapters have been written to be completely independent of one another and thus can be studied in either order, as desired.*** Furthermore, the two chapters have the same structure, making it easier and faster to study the second device, as well as to draw comparisons between the two device types.

Each of Chapters 5 and 6 begins with a study of the device structure and its physical operation, leading to a description of its terminal characteristics. Then, to allow the student to become very familiar with the operation of the transistor as a circuit element, a large number of examples are presented of dc circuits utilizing the device. The last section of each of Chapters 5 and 6 deals with second-order effects that are included for completeness, but that can be skipped if time does not permit detailed coverage.

Chapter 7. The heart of a first course in electronics is the study of transistor amplifiers. Chapter 7 (new to this edition) presents a unified treatment of the subject. It begins with the basic principles that underlie the operation of a transistor, of either type, as an amplifier, and proceeds to present the important concepts of small-signal operation and modeling. This is followed by a study of the basic configurations of single-transistor amplifiers. After a presentation of dc biasing methods, the chapter concludes with practical examples of discrete-circuit amplifiers. The combined presentation emphasizes the unity of the basic principles while allowing for separate treatment of the two device types where this is warranted. Very importantly, we are able to compare the two devices and to draw conclusions about their unique areas of application.

After the study of Part I, the reader will be fully prepared to study either integrated-circuit amplifiers in Part II, or digital integrated circuits in Part III.

Part II, *Integrated-Circuit Amplifiers*, is devoted to the study of practical amplifier circuits that can be fabricated in the integrated-circuit (IC) form. Its six chapters constitute a coherent treatment of IC amplifier design and can thus serve as a second course in electronic circuits.

MOS and Bipolar. Throughout Part II, both MOS and bipolar circuits are presented side-by-side. Because the MOSFET is by far the dominant device, its circuits are presented first. Bipolar circuits are discussed to the same depth but occasionally more briefly.

Chapter 8. Beginning with a brief introduction to the philosophy of IC design, Chapter 8 presents the basic circuit building blocks that are used in the design of IC amplifiers. These include current mirrors, current sources, gain cells, and cascode amplifiers.

Chapter 9. The most important IC building block, the differential pair, is the main topic of Chapter 9. The last section of Chapter 9 is devoted to the study of multistage amplifiers.

Chapter 10. Chapter 10 presents a comprehensive treatment of the important subject of amplifier frequency response. Here, Sections 10.1, 10.2, and 10.3 contain essential material; Section 10.4 provides an in-depth treatment of very useful new tools; and Sections 10.5 to 10.8 present the frequency response analysis of a variety of amplifier configurations that can be studied as and when needed. A selection of the latter sections can be made depending on the time available and the instructor's preference.

Chapter 11. The fourth of the essential topics of Part II, feedback, is the subject of Chapter 11. Both the theory of negative feedback and its application in the design of practical feedback amplifiers are presented. We also discuss the stability problem in feedback amplifiers and treat frequency compensation in some detail.

Chapter 12. In Chapter 12 we switch gears from dealing with small-signal amplifiers to those that are required to handle large signals and large amounts of power. Here we study the different amplifier classes—A, B, and AB—and their realization in bipolar and CMOS technologies. We also consider power BJTs and power MOSFETs, and study representative IC power amplifiers. A brief study of the increasingly popular Class D amplifier is also presented. Depending on the availability of time, some of the later sections can be skipped in a first reading.

Chapter 13. Finally, Chapter 13 brings together all the topics of Part II in an important application; namely, the design of operational amplifier circuits. We study both CMOS and bipolar op amps. In the latter category, besides the classical and still timely 741 circuit, we present modern techniques for the design of low-voltage op amps (Section 13.4).

Part III, *Digital Integrated Circuits*, provides a brief but nonetheless comprehensive and sufficiently detailed study of digital IC design. Our treatment is almost self-contained, requiring for the most part only a thorough understanding of the MOSFET material presented in Chapter 5. Thus, Part III can be studied right after Chapter 5. The only exceptions to this are the last section in Chapter 15 which requires knowledge of the BJT (Chapter 6). Also, knowledge of the MOSFET internal capacitances (Section 10.2.2) will be needed.

Chapter 14. Chapter 14 is the foundation of Part III. It begins with the motivating topic of CMOS logic-gate circuits. Then, following a detailed study of digital logic inverters, we concentrate on the CMOS inverter; its static and dynamic characteristics and its design. Transistor sizing and power dissipation round out the topics of Chapter 14. The material covered in this chapter is the minimum needed to learn something meaningful about digital circuits.

Chapter 15. Chapter 15 has a modular structure and presents six topics of somewhat advanced nature. It begins with a presentation of Moore's law and the technology scaling that has made the multi-billion-transistor chip possible. This is followed by an overview of digital IC technologies, and the design methodologies that make the design of super-complex digital ICs possible. Four different logic-circuit types are then presented. Only the last of these includes bipolar transistors.

Chapter 16. Digital circuits can be broadly divided into logic and memory circuits. The latter is the subject of Chapter 16.

Part IV, *Filters and Oscillators*, is intentionally oriented toward applications and systems. The two topics illustrate powerfully and dramatically the application of both negative and positive feedback.

Chapter 17. Chapter 17 deals with the design of filters, which are important building blocks of communication and instrumentation systems. A comprehensive, design-oriented treatment of the subject is presented. The material provided should allow the reader to perform a complete filter design, starting from specification and ending with a complete circuit realization. A wealth of design tables is included.

Chapter 18. Chapter 18 deals with circuits for the generation of signals with a variety of waveforms: sinusoidal, square, and triangular. We also present circuits for the nonlinear shaping of waveforms.

Appendices. The twelve appendices contain much useful background and supplementary material. We wish to draw the reader's attention in particular to the first two: Appendix A provides a concise introduction to the important topic of IC fabrication technology including IC layout. Appendix B provides SPICE device models as well as a large number of design and simulation examples in PSpice® and Multisim™. The examples are keyed to the book chapters. These Appendices and a great deal more material on these simulation examples can be found on the Companion Website.

Ancillaries

A complete set of ancillary materials is available with this text to support your course.

For the Instructor

The Ancillary Resource Center (ARC) at www.oup-arc.com/sedrasmith is a convenient destination for all the instructor resources that accompany *Microelectronic Circuits*. Accessed online through individual user accounts, the ARC provides instructors with access to up-to-date ancillaries at any time while guaranteeing the security of grade-significant resources. The ARC replaces the Instructor's Resource CD that accompanied the sixth edition. On the ARC, you will find:

- **An electronic version of the Instructor's Solutions Manual.**
- **PowerPoint-based figure slides** that feature all the images and summary tables from the text, with their captions, so they can easily be displayed and explained in class.
- Detailed **instructor's support** for the SPICE circuit simulations in Multisim™ and PSpice®.

The **Instructor's Solutions Manual** (ISBN 978-0-19-933915-0), written by Adel Sedra, contains detailed solutions to all in-text exercises and end-of-chapter problems found in *Microelectronic Circuits*. The Instructor's Solutions Manual for ***Laboratory Explorations to Accompany Microelectronic Circuits*** (ISBN 978-0-19-933926-6) contains detailed solutions to all the exercises and problems found in this student's laboratory guide.

For the Student and Instructor

A Companion Website at www.oup.com/us/sedrasmith features permanently cached versions of device datasheets, so students can design their own circuits in class. The website also contains SPICE circuit simulation examples and lessons. Bonus text topics and the Appendices are also featured on the website.

The ***Laboratory Explorations to Accompany Microelectronic Circuits*** (ISBN 978-0-19-933925-9) invites students to explore the realm of real-world engineering through practical, hands-on experiments. Keyed to sections in the text and taking a "learn-by-doing" approach, it presents labs that focus on the development of practical engineering skills and design practices.

Acknowledgments

Many of the changes in this seventh edition were made in response to feedback received from instructors who adopted the sixth edition. We are grateful to all those who took the time to write to us. In addition, many of the reviewers provided detailed commentary on the sixth edition and suggested a number of the changes that we have incorporated in this edition. They are listed later; to all of them, we extend our sincere thanks. Adel Sedra is also grateful for the feedback received from the students who have taken his electronics courses over the past number of years at the University of Waterloo.

A number of individuals made significant contributions to this edition. Vincent Gaudet of the University of Waterloo contributed to Part III as well as co-authoring the laboratory manual. Wai-Tung Ng of the University of Toronto contributed to Chapter 12 and updated Appendix A (of which he is the original author). Muhammad Faisal of the University of Michigan updated Appendix B, which he helped create for the sixth edition; helped in obtaining the cover photo, and has over a number of years been the source of many good ideas. Olivier Trescases and his students at the University of Toronto pioneered the laboratory system described elsewhere in the Preface. Jennifer Rodrigues typed all the revisions, as she did for a number of the previous editions, with tremendous skill and good humour. Chris Schroeder was of great assistance to Adel Sedra with local logistics. Laura Fujino assisted in many ways and in particular with the "Expand-Your-Perspective" notes. To all of these friends and colleagues we say thank you.

Over the recent years we have benefited greatly from discussions with a number of colleagues and friends. In particular we are very grateful to the following: James Barby, University of Waterloo; David Nairn, University of Waterloo; Anthony Chan Carusone, University of Toronto; David Johns, University of Toronto; Ken Martin, University of Toronto; Khoman Phang, University of Toronto; Gordon Roberts, McGill University; Ali Sheikholeslami, University of Toronto; and Amir Yazdani, Ryerson University.

The cover photograph shows a 3D IC system, which demonstrates the concept of wireless power delivery and communication through multiple layers of CMOS chips. The communication circuits were demonstrated in an IBM 45 nm SOI CMOS process. This technology is designed to serve a multi-Gb/s interconnect between cores spread across several IC layers for high-performance processors. We are grateful to Professor David Wentzloff, Director of the Wireless Integrated Circuits Group at the University of Michigan, who allowed us to use this image, and to Muhammad Faisal, Founder of Movellus Circuits Incorporated, who edited the image.

A large number of people at Oxford University Press contributed to the development of this edition and its various ancillaries. We would like to specifically mention Marketing Manager David Jurman, Marketing Director Frank Mortimer, Higher Ed Sales Director Bill Marting, Copywriter Kristin Maffei, Art Director Michele Laseau, Production Manager Lisa Grzan, Team Leader Amy Whitmer, and Senior Production Editor Jane Lee.

We wish to extend special thanks to our Publisher at Oxford University Press, John Challice, and the Editorial Director, Patrick Lynch. Both have always shown great interest in this book and have provided considerable guidance and support throughout the process of preparing this edition. The Senior Acquisitions Editor, Nancy Blaine, and Associate Editor, Christine Mahon, have done a truly outstanding job. It has been a pleasure to work with both of them, both as professionals and as highly thoughtful individuals; we owe them much gratitude. On the production side, Barbara Mathieu, Senior Production Editor, has been superb: her attention to detail and emphasis on quality is without par.

Finally, we wish to thank our families for their support and understanding, and to thank all the students and instructors who have valued this book throughout its history.

Adel S. Sedra
Kenneth C. (KC) Smith
Waterloo, Ontario, Canada
August 2014

Reviewers of Seventh Edition

Junseok Chae, Arizona State University, Tempe, AZ
Liang Dong, Baylor University, Waco, TX
Muhammad Faisal, University of Michigan, Ann Arbor, MI
Patrick Fay, University of Notre Dame, Notre Dame, IN
Vincent Gaudet, University of Waterloo, Waterloo, Canada
Elmer A Grubbs, Northern Arizona University, Flagstaff, AZ
Serhiy Levkov, New Jersey Institute of Technology, Newark, NJ
Leda Lunardi, North Carolina State University, Raleigh, NC
Phyllis R. Nelson, California State Polytechnic University, Pomona, CA
Robert W. Newcomb, University of Maryland, College Park, MD
Toshikazu Nishida, University of Florida, Gainesville, FL
Matthew Swabey, Purdue University, West Lafayette, IN
Khalid Hasan Tantawi, University of Alabama, Huntsville, AL
Farid M. Tranjan, University of North Carolina, Charlotte, NC
Mustapha C.E. Yagoub, University of Ottawa, Ottawa, Canada
Justin Jackson, Weber State University, Ogden, UT
John Mankowski, Texas Tech University, Lubbock, TX
Chris Mi, University of Michigan, Dearborn, MI

Reviewers of Prior Editions

Maurice Aburdene, Bucknell University, Lewisburg, PA
Michael Bartz, University of Memphis, TN
Elizabeth Brauer, Northern Arizona University, Flagstaff, AZ
Martin Brooke, Duke University, Durham, NC
Patrick L. Chapman, University of Illinois, Urbana–Champaign, IL
Yun Chiu, University of Illinois, Urbana–Champaign, IL
Roy H. Cornely, New Jersey Institute of Technology, Newark, NJ
Norman Cox, Missouri University of Science and Technology, Rolla, MO
Dale L. Critchlow, University of Vermont, Burlington, VT
Robert Bruce Darling, University of Washington, Seattle, WA
Artice Davis, San Jose State University, CA
John Davis, University of Texas, Austin, TX
Christopher DeMarco, University of Wisconsin, Madison, WI
Robert Engelken, Arkansas State University, Jonesboro, AR
Ethan Farquhar, University of Tennessee, Knoxville, TN
Eby G. Friedman, University of Rochester, NY
Paul M. Furth, New Mexico State University, Las Cruces, NM
Rhett T. George, Jr., Duke University, Durham, NC
Roobik Gharabagi, St. Louis University, MO
George Giakos, University of Akron, OH
John Gilmer, Wilkes University, Wilkes-Barre, PA
Michael Green, University of California, Irvine, CA
Steven de Haas, California State University, Sacramento, CA
Anas Hamoui, McGill University, Montreal, Canada
Reza Hashemian, Northern Illinois University, DeKalb, IL
William Harrell, Clemson University, SC
Reid Harrison, University of Utah, Salt Lake City, UT
Ward J. Helms, University of Washington, Seattle, WA
Richard Hornsey, York University, Toronto, Canada
Timothy Horiuchi, University of Maryland, College Park, MD
Hsiung Hsu, The Ohio State University, Columbus, OH
Robert Irvine, California State Polytechnic University, Pomona, CA
Mohammed Ismail, The Ohio State University, Columbus, OH
Paul Israelsen, Utah State University, Logan UT
Steve Jantzi, Broadcom, CA
Zhenhua Jiang, University of Miami, FL
Marian Kazimierczuk, Wright State University, Dayton, OH
John Khoury, Columbia University, New York, NY
Jacob B. Khurgin, The Johns Hopkins University, Baltimore, MD
Seongsin M. Kim, University of Alabama, Tuscaloosa, AL
Roger King, University of Toledo, OH
Clark Kinnaird, Southern Methodist University, Dallas, TX
Robert J. Krueger, University of Wisconsin, Milwaukee, WI
Joy Laskar, Georgia Institute of Technology, Atlanta, GA
Tsu-Jae King Liu, University of California, Berkeley, CA
Yicheng Lu, Rutgers University, Piscataway, NJ
David Luke, University of New Brunswick, Fredericton, Canada
Thomas Matthews, California State University, Sacramento, CA
Un-Ku Moon, Oregon State University, Corvallis, OR
Bahram Nabet, Drexel University, Philadelphia, PA
Dipankar Nagchoudhuri, Indian Institute of Technology, Delhi, India
David Nairn, University of Waterloo, Waterloo, Canada
Joseph H. Nevin, University of Cincinnati, OH
Ken Noren, University of Idaho, Moscow, ID
Brita Olson, California Polytechnic University, Pomona, CA
Martin Peckerar, University of Maryland, College Park, MD
Khoman Phang, University of Toronto, Canada
Mahmudur Rahman, Santa Clara University, CA
Rabin Raut, Concordia University, Montreal, Canada
John A. Ringo, Washington State University, Pullman, WA
Zvi S. Roth, Florida Atlantic University, Boca Raton, FL
Mulukutla Sarma, Northeastern University, Boston, MA
John Scalzo, Louisiana State University, Baton Rouge, LA
Norman Scheinberg, City College, New York, NY
Pierre Schmidt, Florida International University, Miami, FL
Richard Schreier, Analog Devices, Toronto, Canada
Dipankar Sengupta, Royal Melbourne Institute of Technology, Australia
Ali Sheikholeslami, University of Toronto, Canada
Kuang Sheng, Rutgers University, Piscataway, NJ
Michael L. Simpson, University of Tennessee, Knoxville, TN
Karl A. Spuhl, Washington University in St. Louis, MO
Charles Sullivan, Dartmouth College, Hanover, NH
Andrew Szeto, San Diego State University, CA
Joel Therrien, University of Massachusetts, Lowell, MA
Len Trombetta, University of Houston, TX
Daniel van der Weide, University of Delaware, Newark, DE
Gregory M. Wierzba, Michigan State University, East Lansing, MI
Donna Yu, North Carolina State University, Raleigh, NC
Jiann-Shiun Yuan, University of Central Florida, Orlando, FL
Sandra Yost, University of Detroit, Mercy, MI
Alex Zaslavsky, Brown University, Providence, RI
Jianhua (David) Zhang, University of Illinois, Urbana–Champaign, IL

Microelectronic Circuits

PART I

Devices and Basic Circuits

Part I, *Devices and Basic Circuits*, includes the most fundamental and essential topics for the study of electronic circuits. At the same time, it constitutes a complete package for a first course on the subject.

The heart of Part I is the study of the three basic semiconductor devices: the diode (Chapter 4), the MOS transistor (Chapter 5), and the bipolar transistor (Chapter 6). In each case, we study the device operation, its characterization, and its basic circuit applications. Chapter 7 then follows with a study of the most fundamental application of the two transistor types; namely, their use in amplifier design. This side-by-side study of MOSFET and BJT amplifiers allows us to see similarities between these amplifiers and to compare them, which in turn highlights the distinct areas of applicability of each, as well as showing the unity of the basic principles that underlie the use of transistors as amplifiers.

For those who have not had a prior course on device physics, Chapter 3 provides an overview of semiconductor concepts at a level sufficient for the study of electronic circuits. A review of Chapter 3 should prove useful even for those with prior knowledge of semiconductors.

Since the purpose of electronic circuits is the processing of signals, it is essential to understand signals, their characterization in the time and frequency domains, and their analog and digital representations. The basis for such understanding is provided in Chapter 1, which also introduces the most common signal-processing function, *amplification*, and the characterization and types of *amplifiers*.

Besides diodes and transistors, the basic electronic devices, the op amp is studied in Part I. Although not an electronic device in the most fundamental sense, the op amp is commercially available as an integrated circuit (IC) package and has well-defined terminal characteristics. Thus, even though the op amp's internal circuit is complex, typically incorporating 20 or more transistors, its almost-ideal terminal behavior makes it possible to treat the op amp as a circuit element and to use it in the design of powerful circuits, as we do in Chapter 2, without any knowledge of its internal construction. We should mention, however, that the study of op amps can be delayed until a later point, and Chapter 2 can be skipped with no loss of continuity.

The foundation of this book, and of any electronics course, is the study of the two transistor types in use today: the MOS transistor in Chapter 5 and the bipolar transistor in Chapter 6. These two chapters have been written to be completely independent of each other and thus can be studied in either order, as desired.

After the study of Part I, the reader will be fully prepared to undertake the study of either integrated-circuit amplifiers in Part II or digital integrated circuits in Part III.

CHAPTER 1

Signals and Amplifiers

IN THIS CHAPTER YOU WILL LEARN

1. That electronic circuits process signals, and thus understanding electrical signals is essential to appreciating the material in this book.
2. The Thévenin and Norton representations of signal sources.
3. The representation of a signal as the sum of sine waves.
4. The analog and digital representations of a signal.
5. The most basic and pervasive signal-processing function: signal amplification, and correspondingly, the signal amplifier.
6. How amplifiers are characterized (modeled) as circuit building blocks independent of their internal circuitry.
7. How the frequency response of an amplifier is measured, and how it is calculated, especially in the simple but common case of a single-time-constant (STC) type response.

Introduction

The subject of this book is modern electronics, a field that has come to be known as **microelectronics**. **Microelectronics** refers to the integrated-circuit (IC) technology that at the time of this writing is capable of producing circuits that contain billions of components in a small piece of silicon (known as a **silicon chip**) whose area is on the order of 100 mm^2. One such microelectronic circuit, for example, is a complete digital computer, which accordingly is known as a **microcomputer** or, more generally, a **microprocessor**. The microelectronic circuits you will learn to design in this book are used in almost every device we encounter in our daily lives: in the appliances we use in our homes; in the vehicles and transportation systems we use to travel; in the cell phones we use to communicate; in the medical equipment we need to care for our health; in the computers we use to do our work; and in the audio and video systems, the radio and TV sets, and the multitude of other digital devices we use to entertain ourselves. Indeed, it is difficult to conceive of modern life without microelectronic circuits.

In this book we shall study electronic devices that can be used singly (in the design of **discrete circuits**) or as components of an **integrated-circuit (IC)** chip. We shall study the design and analysis of interconnections of these devices, which form discrete and integrated

circuits of varying complexity and perform a wide variety of functions. We shall also learn about available IC chips and their application in the design of electronic systems.

The purpose of this first chapter is to introduce some basic concepts and terminology. In particular, we shall learn about signals and about one of the most important signal-processing functions electronic circuits are designed to perform, namely, signal amplification. We shall then look at circuit representations or models for linear amplifiers. These models will be employed in subsequent chapters in the design and analysis of actual amplifier circuits.

In addition to motivating the study of electronics, this chapter serves as a bridge between the study of linear circuits and that of the subject of this book: the design and analysis of electronic circuits.

1.1 Signals

Signals contain information about a variety of things and activities in our physical world. Examples abound: Information about the weather is contained in signals that represent the air temperature, pressure, wind speed, etc. The voice of a radio announcer reading the news into a microphone provides an acoustic signal that contains information about world affairs. To monitor the status of a nuclear reactor, instruments are used to measure a multitude of relevant parameters, each instrument producing a signal.

To extract required information from a set of signals, the observer (be it a human or a machine) invariably needs to **process** the signals in some predetermined manner. This **signal processing** is usually most conveniently performed by electronic systems. For this to be possible, however, the signal must first be converted into an electrical signal, that is, a voltage or a current. This process is accomplished by devices known as **transducers**. A variety of transducers exist, each suitable for one of the various forms of physical signals. For instance, the sound waves generated by a human can be converted into electrical signals by using a microphone, which is in effect a pressure transducer. It is not our purpose here to study transducers; rather, we shall assume that the signals of interest already exist in the electrical domain and represent them by one of the two equivalent forms shown in Fig. 1.1. In Fig. 1.1(a) the signal is represented by a voltage source $v_s(t)$ having a source resistance R_s. In the alternate representation of Fig. 1.1(b) the signal is represented by a current source $i_s(t)$ having a source resistance R_s. Although the two representations are equivalent, that in Fig. 1.1(a) (known as the Thévenin form) is preferred when R_s is low. The representation of Fig. 1.1(b) (known as the Norton form) is preferred when R_s is high. The reader will come to appreciate this point later in this chapter when we study the different types of amplifiers. For the time being, it is important to be familiar with Thévenin's and Norton's theorems (for a

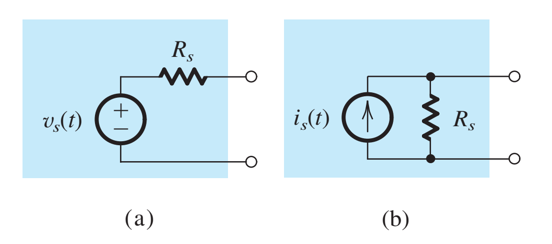


Figure 1.1 Two alternative representations of a signal source: **(a)** the Thévenin form; **(b)** the Norton form.

brief review, see Appendix D) and to note that for the two representations in Fig. 1.1 to be equivalent, their parameters are related by

$$v_s(t) = R_s i_s(t)$$

Example 1.1

The output resistance of a signal source, although inevitable, is an imperfection that limits the ability of the source to deliver its full signal strength to a **load**. To see this point more clearly, consider the signal source when connected to a load resistance R_L as shown in Fig. 1.2. For the case in which the source is represented by its Thévenin equivalent form, find the voltage v_o that appears across R_L, and hence the condition that R_s must satisfy for v_o to be close to the value of v_s. Repeat for the Norton-represented source; in this case finding the current i_o that flows through R_L and hence the condition that R_s must satisfy for i_o to be close to the value of i_s.

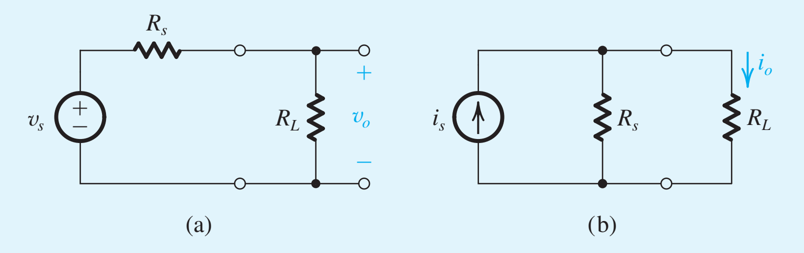


Figure 1.2 Circuits for Example 1.1.

Solution

For the Thévenin-represented signal source shown in Fig. 1.2(a), the output voltage v_o that appears across the load resistance R_L can be found from the ratio of the voltage divider formed by R_s and R_L,

$$v_o = v_s \frac{R_L}{R_L + R_s}$$

From this equation we see that for

$$v_o \simeq v_s$$

the source resistance R_s must be much lower than the load resistance R_L,

$$R_s \ll R_L$$

Thus, for a source represented by its Thévenin equivalent, ideally $R_s = 0$, and as R_s is increased, relative to the load resistance R_L with which this source is intended to operate, the voltage v_o that appears across the load becomes smaller, not a desirable outcome.

Example 1.1 *continued*

Next, we consider the Norton-represented signal source in Fig. 1.2(b). To obtain the current i_o that flows through the load resistance R_L, we utilize the ratio of the current divider formed by R_s and R_L,

$$i_o = i_s \frac{R_s}{R_s + R_L}$$

From this relationship we see that for

$$i_o \simeq i_s$$

the source resistance R_s must be much larger than R_L,

$$R_s \gg R_L$$

Thus for a signal source represented by its Norton equivalent, ideally $R_s = \infty$, and as R_s is reduced, relative to the load resistance R_L with which this source is intended to operate, the current i_o that flows through the load becomes smaller, not a desirable outcome.

Finally, we note that although circuit designers cannot usually do much about the value of R_s, they may have to devise a circuit solution that minimizes or eliminates the loss of signal strength that results when the source is connected to the load.

EXERCISES

1.1 For the signal-source representations shown in Figs. 1.1(a) and 1.1(b), what are the open-circuit output voltages that would be observed? If, for each, the output terminals are short-circuited (i.e., wired together), what current would flow? For the representations to be equivalent, what must the relationship be between v_s, i_s, and R_s?
Ans. For (a), $v_{oc} = v_s(t)$; for (b), $v_{oc} = R_s i_s(t)$; for (a), $i_{sc} = v_s(t)/R_s$; for (b), $i_{sc} = i_s(t)$; for equivalency, $v_s(t) = R_s i_s(t)$

1.2 A signal source has an open-circuit voltage of 10 mV and a short-circuit current of 10 μA. What is the source resistance?
Ans. 1 kΩ

1.3 A signal source that is most conveniently represented by its Thévenin equivalent has $v_s = 10$ mV and $R_s = 1$ kΩ. If the source feeds a load resistance R_L, find the voltage v_o that appears across the load for $R_L = 100$ kΩ, 10 kΩ, 1 kΩ, and 100 Ω. Also, find the lowest permissible value of R_L for which the output voltage is at least 80% of the source voltage.
Ans. 9.9 mV; 9.1 mV; 5 mV; 0.9 mV; 4 kΩ

1.4 A signal source that is most conveniently represented by its Norton equivalent form has $i_s = 10$ μA and $R_s = 100$ kΩ. If the source feeds a load resistance R_L, find the current i_o that flows through the load for $R_L = 1$ kΩ, 10 kΩ, 100 kΩ, and 1 MΩ. Also, find the largest permissible value of R_L for which the load current is at least 80% of the source current.
Ans. 9.9 μA; 9.1 μA; 5 μA; 0.9 μA; 25 kΩ

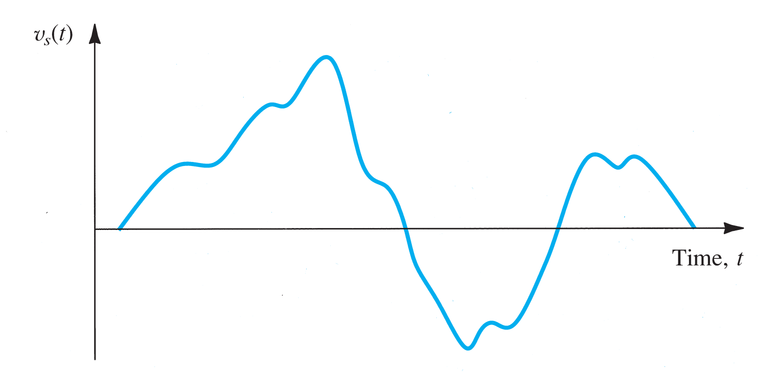


Figure 1.3 An arbitrary voltage signal $v_s(t)$.

From the discussion above, it should be apparent that a signal is a time-varying quantity that can be represented by a graph such as that shown in Fig. 1.3. In fact, the information content of the signal is represented by the changes in its magnitude as time progresses; that is, the information is contained in the "wiggles" in the signal waveform. In general, such waveforms are difficult to characterize mathematically. In other words, it is not easy to describe succinctly an arbitrary-looking waveform such as that of Fig. 1.3. Of course, such a description is of great importance for the purpose of designing appropriate signal-processing circuits that perform desired functions on the given signal. An effective approach to signal characterization is studied in the next section.

1.2 Frequency Spectrum of Signals

An extremely useful characterization of a signal, and for that matter of any arbitrary function of time, is in terms of its **frequency spectrum**. Such a description of signals is obtained through the mathematical tools of **Fourier series** and **Fourier transform**.[1] We are not interested here in the details of these transformations; suffice it to say that they provide the means for representing a voltage signal $v_s(t)$ or a current signal $i_s(t)$ as the sum of sine-wave signals of different frequencies and amplitudes. This makes the sine wave a very important signal in the analysis, design, and testing of electronic circuits. Therefore, we shall briefly review the properties of the sinusoid.

Figure 1.4 shows a sine-wave voltage signal $v_a(t)$,

$$v_a(t) = V_a \sin \omega t \tag{1.1}$$

where V_a denotes the peak value or amplitude in volts and ω denotes the angular frequency in radians per second; that is, $\omega = 2\pi f$ rad/s, where f is the frequency in hertz, $f = 1/T$ Hz, and T is the period in seconds.

The sine-wave signal is completely characterized by its peak value V_a, its frequency ω, and its phase with respect to an arbitrary reference time. In the case depicted in Fig. 1.4, the time

[1]The reader who has not yet studied these topics should not be alarmed. No detailed application of this material will be made until Chapter 10. Nevertheless, a general understanding of Section 1.2 should be very helpful in studying early parts of this book.

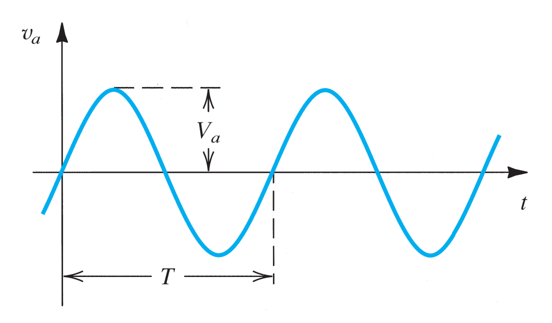


Figure 1.4 Sine-wave voltage signal of amplitude V_a and frequency $f = 1/T$ Hz. The angular frequency $\omega = 2\pi f$ rad/s.

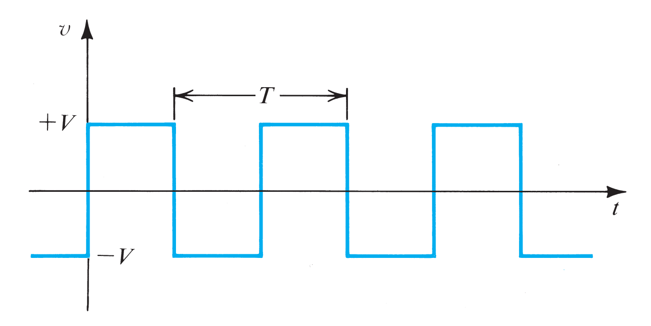


Figure 1.5 A symmetrical square-wave signal of amplitude V.

origin has been chosen so that the phase angle is 0. It should be mentioned that it is common to express the amplitude of a sine-wave signal in terms of its root-mean-square (rms) value, which is equal to the peak value divided by $\sqrt{2}$. Thus the rms value of the sinusoid $v_a(t)$ of Fig. 1.4 is $V_a/\sqrt{2}$. For instance, when we speak of the wall power supply in our homes as being 120 V, we mean that it has a sine waveform of $120\sqrt{2}$ volts peak value.

Returning now to the representation of signals as the sum of sinusoids, we note that the Fourier series is utilized to accomplish this task for the special case of a signal that is a periodic function of time. On the other hand, the Fourier transform is more general and can be used to obtain the frequency spectrum of a signal whose waveform is an arbitrary function of time.

The Fourier series allows us to express a given periodic function of time as the sum of an infinite number of sinusoids whose frequencies are harmonically related. For instance, the symmetrical square-wave signal in Fig. 1.5 can be expressed as

$$v(t) = \frac{4V}{\pi}\left(\sin\omega_0 t + \frac{1}{3}\sin 3\omega_0 t + \frac{1}{5}\sin 5\omega_0 t + \cdots\right) \tag{1.2}$$

where V is the amplitude of the square wave and $\omega_0 = 2\pi/T$ (T is the period of the square wave) is called the **fundamental frequency**. Note that because the amplitudes of the harmonics progressively decrease, the infinite series can be truncated, with the truncated series providing an approximation to the square waveform.

The sinusoidal components in the series of Eq. (1.2) constitute the frequency spectrum of the square-wave signal. Such a spectrum can be graphically represented as in Fig. 1.6, where the horizontal axis represents the angular frequency ω in radians per second.

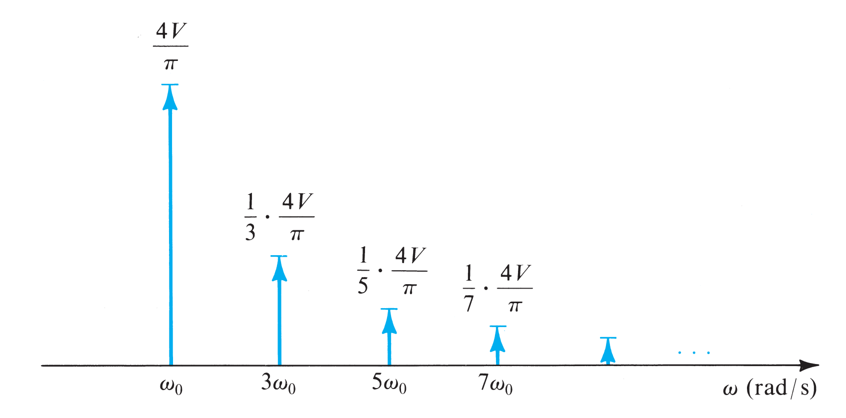


Figure 1.6 The frequency spectrum (also known as the **line spectrum**) of the periodic square wave of Fig. 1.5.

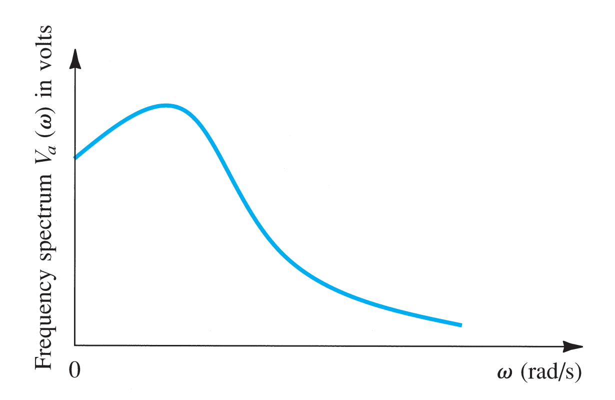


Figure 1.7 The frequency spectrum of an arbitrary waveform such as that in Fig. 1.3.

The Fourier transform can be applied to a nonperiodic function of time, such as that depicted in Fig. 1.3, and provides its frequency spectrum as a continuous function of frequency, as indicated in Fig. 1.7. Unlike the case of periodic signals, where the spectrum consists of discrete frequencies (at ω_0 and its harmonics), the spectrum of a nonperiodic signal contains in general all possible frequencies. Nevertheless, the essential parts of the spectra of practical signals are usually confined to relatively short segments of the frequency (ω) axis—an observation that is very useful in the processing of such signals. For instance, the spectrum of audible sounds such as speech and music extends from about 20 Hz to about 20 kHz—a frequency range known as the **audio band**. Here we should note that although some musical tones have frequencies above 20 kHz, the human ear is incapable of hearing frequencies that are much above 20 kHz. As another example, analog video signals have their spectra in the range of 0 MHz to 4.5 MHz.

We conclude this section by noting that a signal can be represented either by the manner in which its waveform varies with time, as for the voltage signal $v_a(t)$ shown in Fig. 1.3, or in terms of its frequency spectrum, as in Fig. 1.7. The two alternative representations are known as the time-domain representation and the frequency-domain representation, respectively. The frequency-domain representation of $v_a(t)$ will be denoted by the symbol $V_a(\omega)$.

EXERCISES

1.5 Find the frequencies f and ω of a sine-wave signal with a period of 1 ms.
Ans. $f = 1000$ Hz; $\omega = 2\pi \times 10^3$ rad/s

1.6 What is the period T of sine waveforms characterized by frequencies of (a) $f = 60$ Hz? (b) $f = 10^{-3}$ Hz? (c) $f = 1$ MHz?
Ans. 16.7 ms; 1000 s; 1 μs

1.7 The UHF (ultra high frequency) television broadcast band begins with channel 14 and extends from 470 MHz to 806 MHz. If 6 MHz is allocated for each channel, how many channels can this band accommodate?
Ans. 56; channels 14 to 69

1.8 When the square-wave signal of Fig. 1.5, whose Fourier series is given in Eq. (1.2), is applied to a resistor, the total power dissipated may be calculated directly using the relationship $P = 1/T \int_0^T (v^2/R)\,dt$ or indirectly by summing the contribution of each of the harmonic components, that is, $P = P_1 + P_3 + P_5 + \ldots$, which may be found directly from rms values. Verify that the two approaches are equivalent. What fraction of the energy of a square wave is in its fundamental? In its first five harmonics? In its first seven? First nine? In what number of harmonics is 90% of the energy? (Note that in counting harmonics, the fundamental at ω_0 is the first, the one at $2\omega_0$ is the second, etc.)
Ans. 0.81; 0.93; 0.95; 0.96; 3

1.3 Analog and Digital Signals

The voltage signal depicted in Fig. 1.3 is called an **analog signal**. The name derives from the fact that such a signal is *analogous* to the physical signal that it represents. The magnitude of an analog signal can take on any value; that is, the amplitude of an analog signal exhibits a continuous variation over its range of activity. The vast majority of signals in the world around us are analog. Electronic circuits that process such signals are known as **analog circuits**. A variety of analog circuits will be studied in this book.

An alternative form of signal representation is that of a sequence of numbers, each number representing the signal magnitude at an instant of time. The resulting signal is called a **digital signal**. To see how a signal can be represented in this form—that is, how signals can be converted from analog to digital form—consider Fig. 1.8(a). Here the curve represents a voltage signal, identical to that in Fig. 1.3. At equal intervals along the time axis, we have marked the time instants t_0, t_1, t_2, and so on. At each of these time instants, the magnitude of the signal is measured, a process known as **sampling**. Figure 1.8(b) shows a representation of the signal of Fig. 1.8(a) in terms of its samples. The signal of Fig. 1.8(b) is defined only at the sampling instants; it no longer is a continuous function of time; rather, it is a **discrete-time signal**. However, since the magnitude of each sample can take any value in a continuous range, the signal in Fig. 1.8(b) is still an analog signal.

Now if we represent the magnitude of each of the signal samples in Fig. 1.8(b) by a number having a finite number of digits, then the signal amplitude will no longer be continuous; rather,

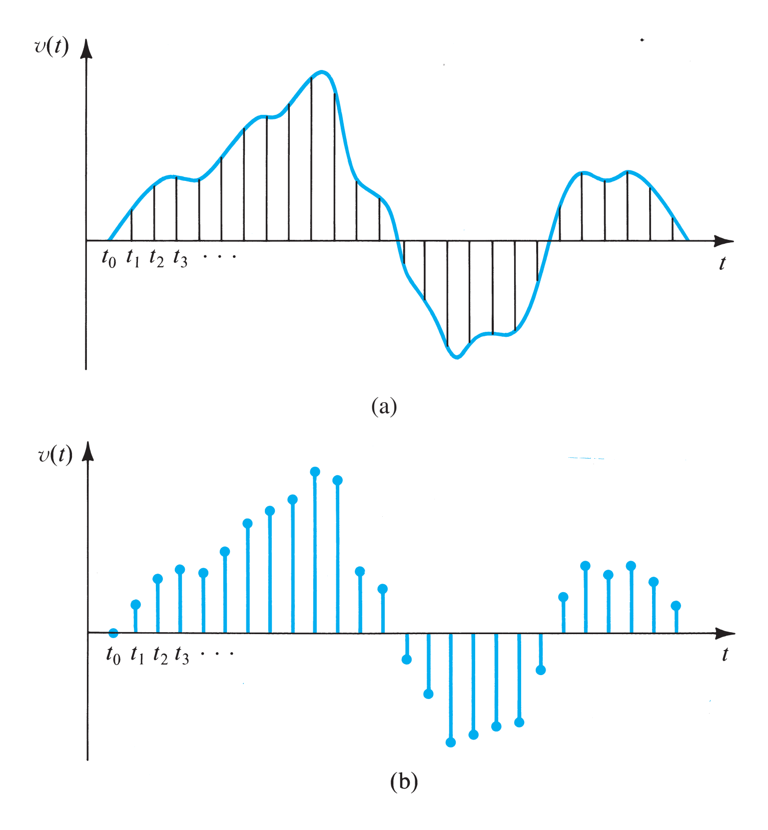


Figure 1.8 Sampling the continuous-time analog signal in **(a)** results in the discrete-time signal in **(b)**.

it is said to be **quantized**, **discretized**, or **digitized**. The resulting digital signal then is simply a sequence of numbers that represent the magnitudes of the successive signal samples.

The choice of number system to represent the signal samples affects the type of digital signal produced and has a profound effect on the complexity of the digital circuits required to process the signals. It turns out that the **binary** number system results in the simplest possible digital signals and circuits. In a binary system, each digit in the number takes on one of only two possible values, denoted 0 and 1. Correspondingly, the digital signals in binary systems need have only two voltage levels, which can be labeled low and high. As an example, in some of the digital circuits studied in this book, the levels are 0 V and +5 V. Figure 1.9 shows the time variation of such a digital signal. Observe that the waveform is a pulse train with 0 V representing a 0 signal, or logic 0, and +5 V representing logic 1.

If we use N *b*inary dig*its* (bits) to represent each sample of the analog signal, then the digitized sample value can be expressed as

$$D = b_0 2^0 + b_1 2^1 + b_2 2^2 + \cdots + b_{N-1} 2^{N-1} \tag{1.3}$$

where $b_0, b_1, \ldots, b_{N-1}$, denote the N bits and have values of 0 or 1. Here bit b_0 is the **least significant bit** (**LSB**), and bit b_{N-1} is the **most significant bit** (**MSB**). Conventionally, this binary number is written as $b_{N-1}\, b_{N-2} \ldots b_0$. We observe that such a representation quantizes the analog sample into one of 2^N levels. Obviously the greater the number of bits (i.e., the larger the N), the closer the digital word D approximates the magnitude of the analog sample. That is, increasing the number of bits reduces the *quantization error* and increases the resolution of the

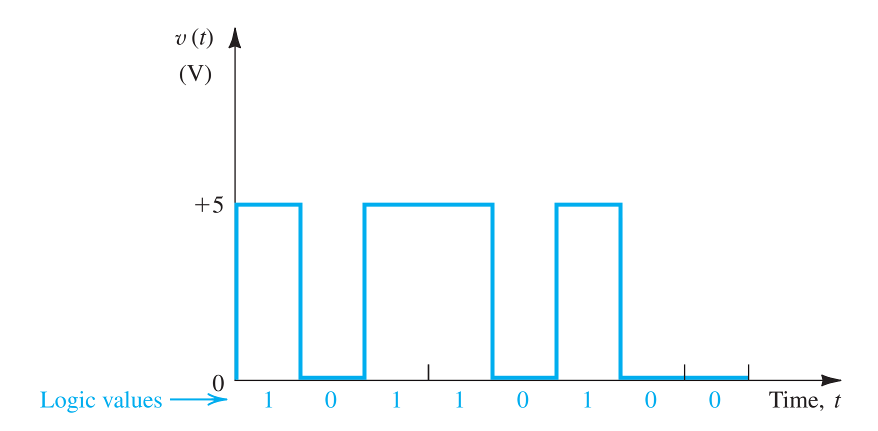


Figure 1.9 Variation of a particular binary digital signal with time.

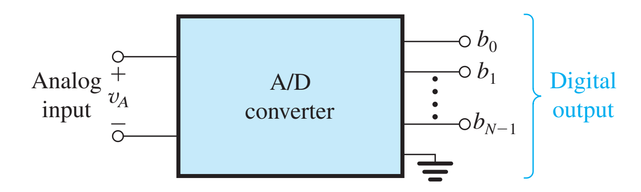


Figure 1.10 Block-diagram representation of the analog-to-digital converter (ADC).

analog-to-digital conversion. This improvement is, however, usually obtained at the expense of more complex and hence more costly circuit implementations. It is not our purpose here to delve into this topic any deeper; we merely want the reader to appreciate the nature of analog and digital signals. Nevertheless, it is an opportune time to introduce a very important circuit building block of modern electronic systems: the **analog-to-digital converter** (**A/D** or **ADC**) shown in block form in Fig. 1.10. The ADC accepts at its input the samples of an analog signal and provides for each input sample the corresponding N-bit digital representation (according to Eq. 1.3) at its N output terminals. Thus although the voltage at the input might be, say, 6.51 V, at each of the output terminals (say, at the ith terminal), the voltage will be either low (0 V) or high (5 V) if b_i is supposed to be 0 or 1, respectively. The dual circuit of the ADC is the **digital-to-analog converter** (**D/A** or **DAC**). It converts an N-bit digital input to an analog output voltage.

Once the signal is in digital form, it can be processed using **digital circuits**. Of course digital circuits can deal also with signals that do not have an analog origin, such as the signals that represent the various instructions of a digital computer.

Since digital circuits deal exclusively with binary signals, their design is simpler than that of analog circuits. Furthermore, digital systems can be designed using a relatively few different kinds of digital circuit blocks. However, a large number (e.g., hundreds of thousands or even millions) of each of these blocks are usually needed. Thus the design of digital circuits poses its own set of challenges to the designer but provides reliable and economic implementations of a great variety of signal-processing functions, many of which are not possible with analog circuits. At the present time, more and more of the signal-processing functions are being performed digitally. Examples around us abound: from the digital watch and the calculator to digital audio systems, digital cameras, and digital television. Moreover, some long-standing

analog systems such as the telephone communication system are now almost entirely digital. And we should not forget the most important of all digital systems, the digital computer.

The basic building blocks of digital systems are logic circuits and memory circuits. We shall study both in this book, beginning in Chapter 14.

One final remark: Although the digital processing of signals is at present all-pervasive, there remain many signal-processing functions that are best performed by analog circuits. Indeed, many electronic systems include both analog and digital parts. It follows that a good electronics engineer must be proficient in the design of both analog and digital circuits, or **mixed-signal** or **mixed-mode** design as it is currently known. Such is the aim of this book.

EXERCISE

1.9 Consider a 4-bit digital word $D = b_3 b_2 b_1 b_0$ (see Eq. 1.3) used to represent an analog signal v_A that varies between 0 V and +15 V.

(a) Give D corresponding to $v_A = 0$ V, 1 V, 2 V, and 15 V.

(b) What change in v_A causes a change from 0 to 1 in (i) b_0, (ii) b_1, (iii) b_2, and (iv) b_3?

(c) If $v_A = 5.2$ V, what do you expect D to be? What is the resulting error in representation?

Ans. (a) 0000, 0001, 0010, 1111; (b) +1 V, +2 V, +4 V, +8 V; (c) 0101, –4%

ANALOG VS. DIGITAL CIRCUIT ENGINEERS:

As digital became the preferred implementation of more and more signal-processing functions, the need arose for greater numbers of digital circuit design engineers. Yet despite predictions made periodically that the demand for analog circuit design engineers would lessen, this has not been the case. Rather, the demand for analog engineers has, if anything, increased. What is true, however, is that the skill level required of analog engineers has risen. Not only are they asked to design circuits of greater sophistication and tighter specifications, but they also have to do this using technologies that are optimized for digital (and not analog) circuits. This is dictated by economics, as digital usually constitutes the larger part of most systems.

1.4 Amplifiers

In this section, we shall introduce the most fundamental signal-processing function, one that is employed in some form in almost every electronic system, namely, signal amplification. We shall study the amplifier as a circuit building block; that is, we shall consider its external characteristics and leave the design of its internal circuit to later chapters.

1.4.1 Signal Amplification

From a conceptual point of view the simplest signal-processing task is that of **signal amplification**. The need for amplification arises because transducers provide signals that

are said to be "weak," that is, in the microvolt (μV) or millivolt (mV) range and possessing little energy. Such signals are too small for reliable processing, and processing is much easier if the signal magnitude is made larger. The functional block that accomplishes this task is the **signal amplifier**.

It is appropriate at this point to discuss the need for **linearity** in amplifiers. Care must be exercised in the amplification of a signal, so that the information contained in the signal is not changed and no new information is introduced. Thus when we feed the signal shown in Fig. 1.3 to an amplifier, we want the output signal of the amplifier to be an exact replica of that at the input, except of course for having larger magnitude. In other words, the "wiggles" in the output waveform must be identical to those in the input waveform. Any change in waveform is considered to be **distortion** and is obviously undesirable.

An amplifier that preserves the details of the signal waveform is characterized by the relationship

$$v_o(t) = Av_i(t) \tag{1.4}$$

where v_i and v_o are the input and output signals, respectively, and A is a constant representing the magnitude of amplification, known as **amplifier gain**. Equation (1.4) is a linear relationship; hence the amplifier it describes is a **linear amplifier**. It should be easy to see that if the relationship between v_o and v_i contains higher powers of v_i, then the waveform of v_o will no longer be identical to that of v_i. The amplifier is then said to exhibit **nonlinear distortion**.

The amplifiers discussed so far are primarily intended to operate on very small input signals. Their purpose is to make the signal magnitude larger, and therefore they are thought of as **voltage amplifiers**. The **preamplifier** in the home stereo system is an example of a voltage amplifier.

At this time we wish to mention another type of amplifier, namely, the **power amplifier**. Such an amplifier may provide only a modest amount of voltage gain but substantial current gain. Thus while absorbing little power from the input signal source to which it is connected, often a preamplifier, it delivers large amounts of power to its load. An example is found in the power amplifier of the home stereo system, whose purpose is to provide sufficient power to drive the loudspeaker, which is the amplifier load. Here we should note that the loudspeaker is the output transducer of the stereo system; it converts the electric output signal of the system into an acoustic signal. A further appreciation of the need for linearity can be acquired by reflecting on the power amplifier. A linear power amplifier causes both soft and loud music passages to be reproduced without distortion.

1.4.2 Amplifier Circuit Symbol

The signal amplifier is obviously a two-port circuit. Its function is conveniently represented by the circuit symbol of Fig. 1.11(a). This symbol clearly distinguishes the input and output ports and indicates the direction of signal flow. Thus, in subsequent diagrams it will not be necessary to label the two ports "input" and "output." For generality we have shown the amplifier to have two input terminals that are distinct from the two output terminals. A more common situation is illustrated in Fig. 1.11(b), where a common terminal exists between the input and output ports of the amplifier. This common terminal is used as a reference point and is called the **circuit ground**.

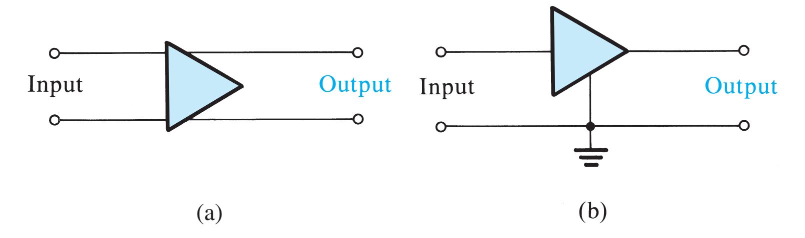


Figure 1.11 **(a)** Circuit symbol for amplifier. **(b)** An amplifier with a common terminal (ground) between the input and output ports.

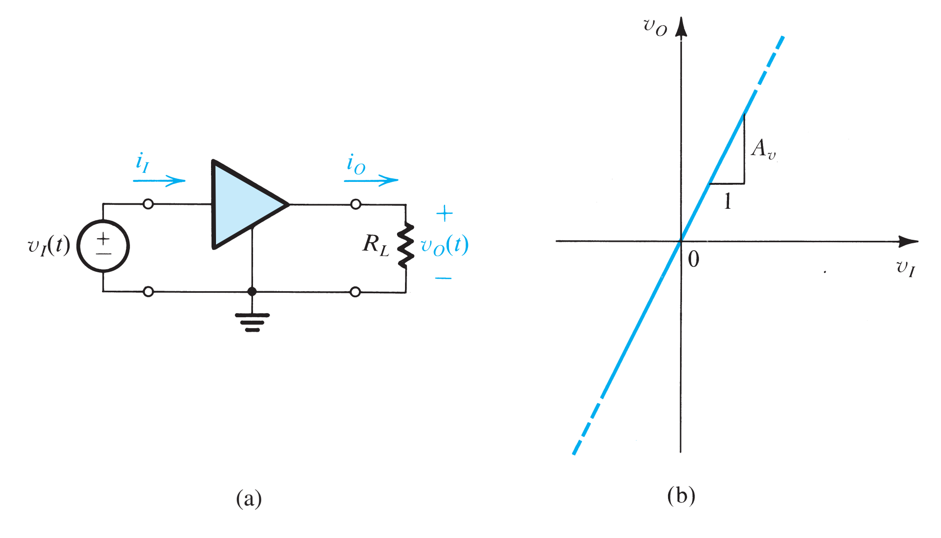


Figure 1.12 **(a)** A voltage amplifier fed with a signal $v_I(t)$ and connected to a load resistance R_L. **(b)** Transfer characteristic of a linear voltage amplifier with voltage gain A_v.

1.4.3 Voltage Gain

A linear amplifier accepts an input signal $v_I(t)$ and provides at the output, across a load resistance R_L (see Fig. 1.12(a)), an output signal $v_O(t)$ that is a magnified replica of $v_I(t)$. The **voltage gain** of the amplifier is defined by

$$\text{Voltage gain } (A_v) = \frac{v_O}{v_I} \tag{1.5}$$

Fig. 1.12(b) shows the **transfer characteristic** of a linear amplifier. If we apply to the input of this amplifier a sinusoidal voltage of amplitude $\hat{V}$, we obtain at the output a sinusoid of amplitude $A_v\hat{V}$.

1.4.4 Power Gain and Current Gain

An amplifier increases the signal power, an important feature that distinguishes an amplifier from a transformer. In the case of a transformer, although the voltage delivered to the load could be greater than the voltage feeding the input side (the primary), the power delivered to the load (from the secondary side of the transformer) is less than or at most equal to the

power supplied by the signal source. On the other hand, an amplifier provides the load with power greater than that obtained from the signal source. That is, amplifiers have power gain. The **power gain** of the amplifier in Fig. 1.12(a) is defined as

$$\text{Power gain } (A_p) \equiv \frac{\text{load power } (P_L)}{\text{input power} (P_I)} \tag{1.6}$$

$$= \frac{v_O i_O}{v_I i_I} \tag{1.7}$$

where i_O is the current that the amplifier delivers to the load $(R_L), i_O = v_O/R_L$, and i_I is the current the amplifier draws from the signal source. The **current gain** of the amplifier is defined as

$$\text{Current gain } (A_i) \equiv \frac{i_O}{i_I} \tag{1.8}$$

From Eqs. (1.5) to (1.8) we note that

$$A_p = A_v A_i \tag{1.9}$$

1.4.5 Expressing Gain in Decibels

The amplifier gains defined above are ratios of similarly dimensioned quantities. Thus they will be expressed either as dimensionless numbers or, for emphasis, as V/V for the voltage gain, A/A for the current gain, and W/W for the power gain. Alternatively, for a number of reasons, some of them historic, electronics engineers express amplifier gain with a logarithmic measure. Specifically the voltage gain A_v can be expressed as

$$\text{Voltage gain in decibels } = 20 \log |A_v| \quad \text{dB}$$

and the current gain A_i can be expressed as

$$\text{Current gain in decibels } = 20 \log |A_i| \quad \text{dB}$$

Since power is related to voltage (or current) squared, the power gain A_p can be expressed in decibels as

$$\text{Power gain in decibels } = 10 \log A_p \quad \text{dB}$$

The absolute values of the voltage and current gains are used because in some cases A_v or A_i will be a negative number. A negative gain A_v simply means that there is a 180° phase difference between input and output signals; it does not imply that the amplifier is **attenuating** the signal. On the other hand, an amplifier whose voltage gain is, say, −20 dB is in fact attenuating the input signal by a factor of 10 (i.e., $A_v = 0.1$ V/V).

1.4.6 The Amplifier Power Supplies

Since the power delivered to the load is greater than the power drawn from the signal source, the question arises as to the source of this additional power. The answer is found by observing that amplifiers need dc power supplies for their operation. These dc sources supply the extra power delivered to the load as well as any power that might be dissipated in the internal circuit

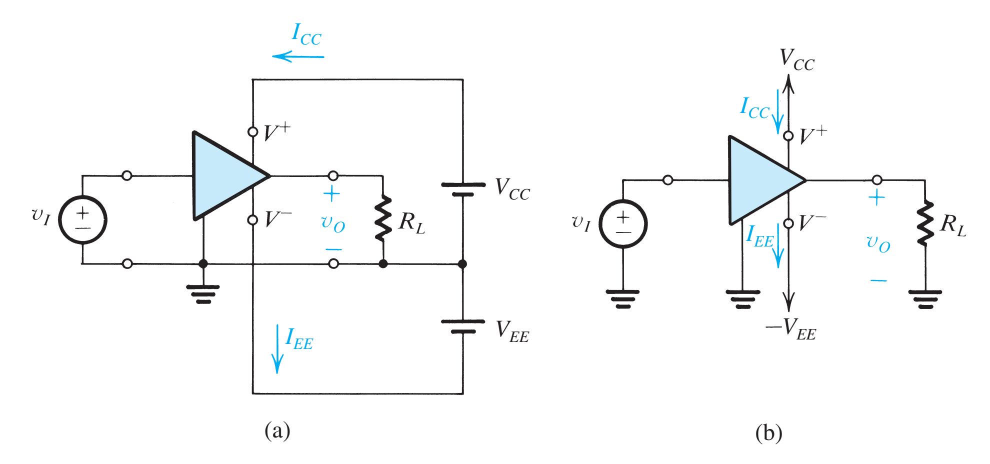


Figure 1.13 An amplifier that requires two dc supplies (shown as batteries) for operation.

of the amplifier (such power is converted to heat). In Fig. 1.12(a) we have not explicitly shown these dc sources.

Figure 1.13(a) shows an amplifier that requires two dc sources: one positive of value V_{CC} and one negative of value V_{EE}. The amplifier has two terminals, labeled V^+ and V^-, for connection to the dc supplies. For the amplifier to operate, the terminal labeled V^+ has to be connected to the positive side of a dc source whose voltage is V_{CC} and whose negative side is connected to the circuit ground. Also, the terminal labeled V^- has to be connected to the negative side of a dc source whose voltage is V_{EE} and whose positive side is connected to the circuit ground. Now, if the current drawn from the positive supply is denoted I_{CC} and that from the negative supply is I_{EE} (see Fig. 1.13a), then the dc power delivered to the amplifier is

$$P_{\text{dc}} = V_{CC}I_{CC} + V_{EE}I_{EE}$$

If the power dissipated in the amplifier circuit is denoted $P_{\text{dissipated}}$, the power-balance equation for the amplifier can be written as

$$P_{\text{dc}} + P_I = P_L + P_{\text{dissipated}}$$

where P_I is the power drawn from the signal source and P_L is the power delivered to the load. Since the power drawn from the signal source is usually small, the amplifier power **efficiency** is defined as

$$\eta \equiv \frac{P_L}{P_{\text{dc}}} \times 100 \qquad (1.10)$$

The power efficiency is an important performance parameter for amplifiers that handle large amounts of power. Such amplifiers, called power amplifiers, are used, for example, as output amplifiers of stereo systems.

In order to simplify circuit diagrams, we shall adopt the convention illustrated in Fig. 1.13(b). Here the V^+ terminal is shown connected to an arrowhead pointing upward and the V^- terminal to an arrowhead pointing downward. The corresponding voltage is indicated next to each arrowhead. Note that in many cases we will not explicitly show the connections

of the amplifier to the dc power sources. Finally, we note that some amplifiers require only one power supply.

Example 1.2

Consider an amplifier operating from ±10-V power supplies. It is fed with a sinusoidal voltage having 1 V peak and delivers a sinusoidal voltage output of 9 V peak to a 1-kΩ load. The amplifier draws a current of 9.5 mA from each of its two power supplies. The input current of the amplifier is found to be sinusoidal with 0.1 mA peak. Find the voltage gain, the current gain, the power gain, the power drawn from the dc supplies, the power dissipated in the amplifier, and the amplifier efficiency.

Solution

$$A_v = \frac{9}{1} = 9 \text{ V/V}$$

or

$$A_v = 20 \log 9 = 19.1 \text{ dB}$$

$$\hat{I}_o = \frac{9 \text{ V}}{1 \text{ k}\Omega} = 9 \text{ mA}$$

$$A_i = \frac{\hat{I}_o}{\hat{I}_i} = \frac{9}{0.1} = 90 \text{ A/A}$$

or

$$A_i = 20 \log 90 = 39.1 \text{ dB}$$

$$P_L = V_{o\text{rms}} I_{o\text{rms}} = \frac{9}{\sqrt{2}} \frac{9}{\sqrt{2}} = 40.5 \text{ mW}$$

$$P_I = V_{i\text{rms}} I_{i\text{rms}} = \frac{1}{\sqrt{2}} \frac{0.1}{\sqrt{2}} = 0.05 \text{ mW}$$

$$A_p = \frac{P_L}{P_I} = \frac{40.5}{0.05} = 810 \text{ W/W}$$

or

$$A_p = 10 \log 810 = 29.1 \text{ dB}$$

$$P_{\text{dc}} = 10 \times 9.5 + 10 \times 9.5 = 190 \text{ mW}$$

$$P_{\text{dissipated}} = P_{\text{dc}} + P_I - P_L$$

$$= 190 + 0.05 - 40.5 = 149.6 \text{ mW}$$

$$\eta = \frac{P_L}{P_{\text{dc}}} \times 100 = 21.3\%$$

From the above example we observe that the amplifier converts some of the dc power it draws from the power supplies to signal power that it delivers to the load.

1.4.7 Amplifier Saturation

Practically speaking, the amplifier transfer characteristic remains linear over only a limited range of input and output voltages. For an amplifier operated from two power supplies the output voltage cannot exceed a specified positive limit and cannot decrease below a specified negative limit. The resulting transfer characteristic is shown in Fig. 1.14, with the positive and negative saturation levels denoted L_+ and L_-, respectively. Each of the two saturation levels is usually within a fraction of a volt of the voltage of the corresponding power supply.

Obviously, in order to avoid distorting the output signal waveform, the input signal swing must be kept within the linear range of operation,

$$\frac{L_-}{A_v} \leq v_I \leq \frac{L_+}{A_v}$$

In Fig. 1.14, which shows two input waveforms and the corresponding output waveforms, the peaks of the larger waveform have been clipped off because of amplifier saturation.

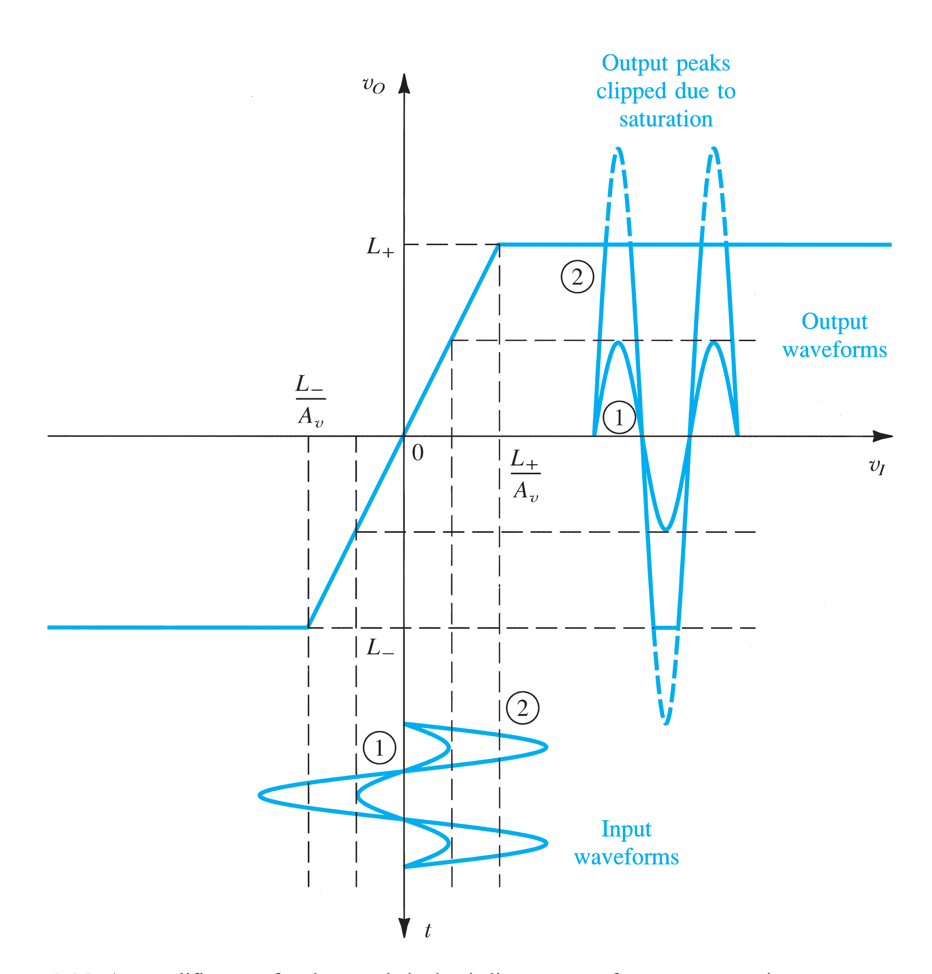


Figure 1.14 An amplifier transfer characteristic that is linear except for output saturation.

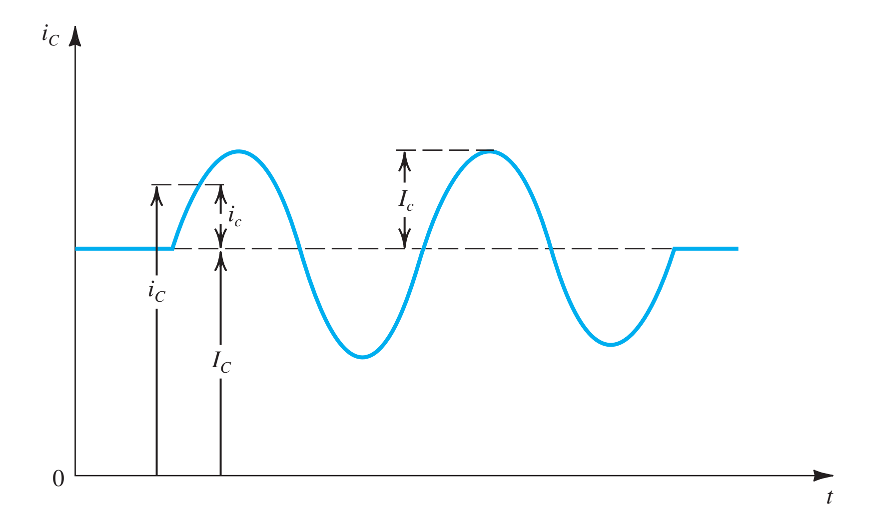


Figure 1.15 Symbol convention employed throughout the book.

1.4.8 Symbol Convention

At this point, we draw the reader's attention to the terminology we shall employ throughout the book. To illustrate the terminology, Fig. 1.15 shows the waveform of a current $i_C(t)$ that is flowing through a branch in a particular circuit. The current $i_C(t)$ consists of a dc component I_C on which is superimposed a sinusoidal component $i_c(t)$ whose peak amplitude is I_c. Observe that at a time t, the **total instantaneous** current $i_C(t)$ is the sum of the dc current I_C and the signal current $i_c(t)$,

$$i_C(t) = I_C + i_c(t) \tag{1.11}$$

where the signal current is given by

$$i_c(t) = I_c \sin \omega t$$

Thus, we state some conventions: Total instantaneous quantities are denoted by a lowercase symbol with uppercase subscript(s), for example, $i_C(t), v_{DS}(t)$. Direct-current (dc) quantities are denoted by an uppercase symbol with uppercase subscript(s), for example, I_C, V_{DS}. Incremental signal quantities are denoted by a lowercase symbol with lowercase subscript(s), for example, $i_c(t), v_{gs}(t)$. If the signal is a sine wave, then its amplitude is denoted by an uppercase symbol with lowercase subscript(s), for example, I_c, V_{gs}. Finally, although not shown in Fig. 1.15, dc power supplies are denoted by an uppercase letter with a double-letter uppercase subscript, for example, V_{CC}, V_{DD}. A similar notation is used for the dc current drawn from the power supply, for example, I_{CC}, I_{DD}.

EXERCISES

1.10 An amplifier has a voltage gain of 100 V/V and a current gain of 1000 A/A. Express the voltage and current gains in decibels and find the power gain.
Ans. 40 dB; 60 dB; 50 dB

1.11 An amplifier operating from a single 15-V supply provides a 12-V peak-to-peak sine-wave signal to a 1-kΩ load and draws negligible input current from the signal source. The dc current drawn from the 15-V supply is 8 mA. What is the power dissipated in the amplifier, and what is the amplifier efficiency?
Ans. 102 mW; 15%

1.5 Circuit Models for Amplifiers

A substantial part of this book is concerned with the design of amplifier circuits that use transistors of various types. Such circuits will vary in complexity from those using a single transistor to those with 20 or more devices. In order to be able to apply the resulting amplifier circuit as a building block in a system, one must be able to characterize, or **model**, its terminal behavior. In this section, we study simple but effective amplifier models. These models apply irrespective of the complexity of the internal circuit of the amplifier. The values of the model parameters can be found either by analyzing the amplifier circuit or by performing measurements at the amplifier terminals.

1.5.1 Voltage Amplifiers

Figure 1.16(a) shows a circuit model for the voltage amplifier. The model consists of a voltage-controlled voltage source having a gain factor A_{vo}, an input resistance R_i that accounts for the fact that the amplifier draws an input current from the signal source, and an output resistance R_o that accounts for the change in output voltage as the amplifier is called upon to supply output current to a load. To be specific, we show in Fig. 1.16(b) the amplifier model fed with a signal voltage source v_s having a resistance R_s and connected at the output to a load resistance R_L. The nonzero output resistance R_o causes only a fraction of $A_{vo}v_i$ to appear across the output. Using the voltage-divider rule we obtain

$$v_o = A_{vo} v_i \frac{R_L}{R_L + R_o}$$

Thus the voltage gain is given by

$$A_v \equiv \frac{v_o}{v_i} = A_{vo} \frac{R_L}{R_L + R_o} \tag{1.12}$$

It follows that in order not to lose gain in coupling the amplifier output to a load, the output resistance R_o should be much smaller than the load resistance R_L. In other words, for a given R_L one must design the amplifier so that its R_o is much smaller than R_L. Furthermore, there are applications in which R_L is known to vary over a certain range. In order to keep the output voltage v_o as constant as possible, the amplifier is designed with R_o much smaller than the lowest value of R_L. An ideal voltage amplifier is one with $R_o = 0$. Equation (1.12) indicates also that for $R_L = \infty, A_v = A_{vo}$. Thus A_{vo} is the voltage gain of the unloaded amplifier, or the **open-circuit voltage gain**. It should also be clear that in specifying the voltage gain of an amplifier, one must also specify the value of load resistance at which this gain is measured or

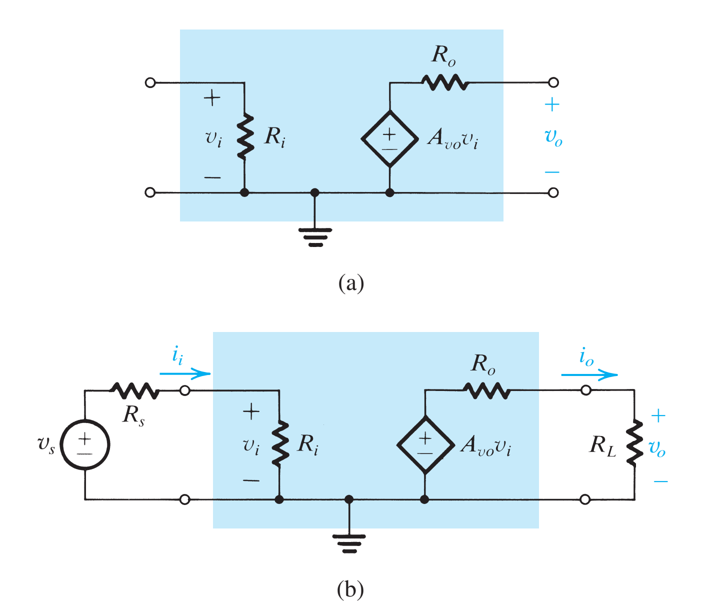


Figure 1.16 **(a)** Circuit model for the voltage amplifier. **(b)** The voltage amplifier with input signal source and load.

calculated. If a load resistance is not specified, it is normally assumed that the given voltage gain is the open-circuit gain A_{vo}.

The finite input resistance R_i introduces another voltage-divider action at the input, with the result that only a fraction of the source signal v_s actually reaches the input terminals of the amplifier; that is,

$$v_i = v_s \frac{R_i}{R_i + R_s} \tag{1.13}$$

It follows that in order not to lose a significant portion of the input signal in coupling the signal source to the amplifier input, the amplifier must be designed to have an input resistance R_i much greater than the resistance of the signal source, $R_i \gg R_s$. Furthermore, there are applications in which the source resistance is known to vary over a certain range. To minimize the effect of this variation on the value of the signal that appears at the input of the amplifier, the design ensures that R_i is much greater than the largest value of R_s. An ideal voltage amplifier is one with $R_i = \infty$. In this ideal case both the current gain and power gain become infinite.

The overall voltage gain (v_o/v_s) can be found by combining Eqs. (1.12) and (1.13),

$$\frac{v_o}{v_s} = A_{vo} \frac{R_i}{R_i + R_s} \frac{R_L}{R_L + R_o}$$

There are situations in which one is interested not in voltage gain but only in a significant power gain. For instance, the source signal can have a respectable voltage but a source resistance that is much greater than the load resistance. Connecting the source directly to the load would result in significant signal attenuation. In such a case, one requires an amplifier with a high input resistance (much greater than the source resistance) and a low output resistance (much smaller than the load resistance) but with a modest voltage gain (or even unity gain).

Such an amplifier is referred to as a **buffer amplifier**. We shall encounter buffer amplifiers often throughout this book.

EXERCISES

1.12 A transducer characterized by a voltage of 1 V rms and a resistance of 1 MΩ is available to drive a 10-Ω load. If connected directly, what voltage and power levels result at the load? If a unity-gain (i.e., $A_{vo} = 1$) buffer amplifier with 1-MΩ input resistance and 10-Ω output resistance is interposed between source and load, what do the output voltage and power levels become? For the new arrangement, find the voltage gain from source to load, and the power gain (both expressed in decibels).
Ans. 10 μV rms; 10^{-11} W; 0.25 V; 6.25 mW; −12 dB; 44 dB

1.13 The output voltage of a voltage amplifier has been found to decrease by 20% when a load resistance of 1 kΩ is connected. What is the value of the amplifier output resistance?
Ans. 250 Ω

1.14 An amplifier with a voltage gain of +40 dB, an input resistance of 10 kΩ, and an output resistance of 1 kΩ is used to drive a 1-kΩ load. What is the value of A_{vo}? Find the value of the power gain in decibels.
Ans. 100 V/V; 44 dB

1.5.2 Cascaded Amplifiers

To meet given amplifier specifications, we often need to design the amplifier as a cascade of two or more stages. The stages are usually not identical; rather, each is designed to serve a specific purpose. For instance, in order to provide the overall amplifier with a large input resistance, the first stage is usually required to have a large input resistance. Also, in order to equip the overall amplifier with a low output resistance, the final stage in the cascade is usually designed to have a low output resistance. To illustrate the analysis and design of cascaded amplifiers, we consider a practical example.

Example 1.3

Figure 1.17 depicts an amplifier composed of a cascade of three stages. The amplifier is fed by a signal source with a source resistance of 100 kΩ and delivers its output into a load resistance of 100 Ω. The first stage has a relatively high input resistance and a modest gain factor of 10. The second stage has a higher gain factor but lower input resistance. Finally, the last, or output, stage has unity gain but a low output resistance. We wish to evaluate the overall voltage gain, that is, v_L/v_s, the current gain, and the power gain.

Example 1.3 *continued*

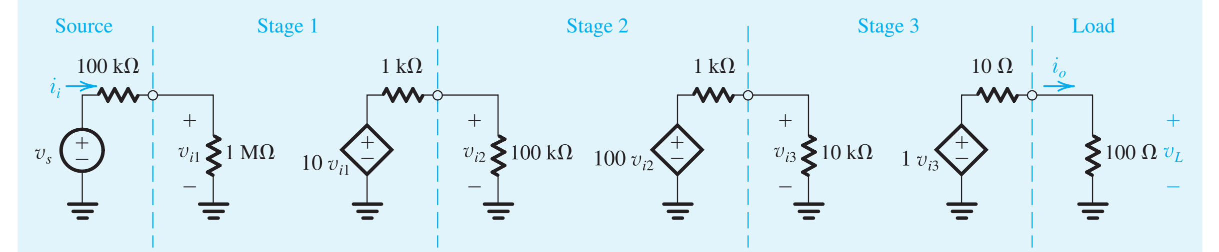


Figure 1.17 Three-stage amplifier for Example 1.3.

Solution

The fraction of source signal appearing at the input terminals of the amplifier is obtained using the voltage-divider rule at the input, as follows:

$$\frac{v_{i1}}{v_s} = \frac{1\text{ M}\Omega}{1\text{ M}\Omega + 100\text{ k}\Omega} = 0.909 \text{ V/V}$$

The voltage gain of the first stage is obtained by considering the input resistance of the second stage to be the load of the first stage; that is,

$$A_{v1} \equiv \frac{v_{i2}}{v_{i1}} = 10\frac{100\text{ k}\Omega}{100\text{ k}\Omega + 1\text{ k}\Omega} = 9.9 \text{ V/V}$$

Similarly, the voltage gain of the second stage is obtained by considering the input resistance of the third stage to be the load of the second stage,

$$A_{v2} \equiv \frac{v_{i3}}{v_{i2}} = 100\frac{10\text{ k}\Omega}{10\text{ k}\Omega + 1\text{ k}\Omega} = 90.9 \text{ V/V}$$

Finally, the voltage gain of the output stage is as follows:

$$A_{v3} \equiv \frac{v_L}{v_{i3}} = 1\frac{100\ \Omega}{100\ \Omega + 10\ \Omega} = 0.909 \text{ V/V}$$

The total gain of the three stages in cascade can now be found from

$$A_v \equiv \frac{v_L}{v_{i1}} = A_{v1}A_{v2}A_{v3} = 818 \text{ V/V}$$

or 58.3 dB.

To find the voltage gain from source to load, we multiply A_v by the factor representing the loss of gain at the input; that is,

$$\frac{v_L}{v_s} = \frac{v_L}{v_{i1}}\frac{v_{i1}}{v_s} = A_v\frac{v_{i1}}{v_s}$$
$$= 818 \times 0.909 = 743.6 \text{ V/V}$$

or 57.4 dB.

The current gain is found as follows:

$$A_i \equiv \frac{i_o}{i_i} = \frac{v_L/100\ \Omega}{v_{i1}/1\ \text{M}\Omega}$$
$$= 10^4 \times A_v = 8.18 \times 10^6 \text{ A/A}$$

or 138.3 dB.

The power gain is found from

$$A_p \equiv \frac{P_L}{P_I} = \frac{v_L i_o}{v_{i1} i_i}$$
$$= A_v A_i = 818 \times 8.18 \times 10^6 = 66.9 \times 10^8 \text{ W/W}$$

or 98.3 dB. Note that

$$A_p(\text{dB}) = \frac{1}{2}[A_v(\text{dB}) + A_i(\text{dB})]$$

A few comments on the cascade amplifier in the above example are in order. To avoid losing signal strength at the amplifier input where the signal is usually very small, the first stage is designed to have a relatively large input resistance (1 MΩ), which is much larger than the source resistance. The trade-off appears to be a moderate voltage gain (10 V/V). The second stage does not need to have such a high input resistance; rather, here we need to realize the bulk of the required voltage gain. The third and final, or output, stage is not asked to provide any voltage gain; rather, it functions as a buffer amplifier, providing a relatively large input resistance and a low output resistance, much lower than R_L. It is this stage that enables connecting the amplifier to the 100-Ω load. These points can be made more concrete by solving the following exercises. In so doing, observe that in finding the gain of an amplifier stage in a cascade amplifier, the loading effect of the succeeding amplifier stage must be taken into account as we have done in the above example.

EXERCISES

1.15 What would the overall voltage gain of the cascade amplifier in Example 1.3 be without stage 3 (i.e., with the load resistance connected to the output of the second stage)?
Ans. 81.8 V/V; a decrease by a factor of 9.

1.16 For the cascade amplifier of Example 1.3, let v_s be 1 mV. Find v_{i1}, v_{i2}, v_{i3}, and v_L.
Ans. 0.91 mV; 9 mV; 818 mV; 744 mV

1.17 (a) Model the three-stage amplifier of Example 1.3 (without the source and load), using the voltage amplifier model of Fig. 1.16(a). What are the values of R_i, A_{vo}, and R_o?
(b) If R_L varies in the range 10 Ω to 1000 Ω, find the corresponding range of the overall voltage gain, v_o/v_s.
Ans. 1 MΩ, 900 V/V, 10 Ω; 409 V/V to 810 V/V

1.5.3 Other Amplifier Types

In the design of an electronic system, the signal of interest—whether at the system input, at an intermediate stage, or at the output—can be either a voltage or a current. For instance, some transducers have very high output resistances and can be more appropriately modeled as current sources. Similarly, there are applications in which the output current rather than the voltage is of interest. Thus, although it is the most popular, the voltage amplifier considered above is just one of four possible amplifier types. The other three are the current amplifier, the transconductance amplifier, and the transresistance amplifier. Table 1.1 shows the four amplifier types, their circuit models, the definition of their gain parameters, and the ideal values of their input and output resistances.

1.5.4 Relationships between the Four Amplifier Models

Although for a given amplifier a particular one of the four models in Table 1.1 is most preferable, *any of the four can be used to model any amplifier*. In fact, simple relationships can be derived to relate the parameters of the various models. For instance, the open-circuit

Table 1.1 The Four Amplifier Types

Type	Circuit Model	Gain Parameter	Ideal Characteristics
Voltage Amplifier		Open-Circuit Voltage Gain $A_{vo} \equiv \left.\frac{v_o}{v_i}\right\|_{i_o=0}$ (V/V)	$R_i = \infty$ $R_o = 0$
Current Amplifier		Short-Circuit Current Gain $A_{is} \equiv \left.\frac{i_o}{i_i}\right\|_{v_o=0}$ (A/A)	$R_i = 0$ $R_o = \infty$
Transconductance Amplifier		Short-Circuit Transconductance $G_m \equiv \left.\frac{i_o}{v_i}\right\|_{v_o=0}$ (A/V)	$R_i = \infty$ $R_o = \infty$
Transresistance Amplifier		Open-Circuit Transresistance $R_m \equiv \left.\frac{v_o}{i_i}\right\|_{i_o=0}$ (V/A)	$R_i = 0$ $R_o = 0$

voltage gain A_{vo} can be related to the short-circuit current gain A_{is} as follows: The open-circuit output voltage given by the voltage amplifier model of Table 1.1 is $A_{vo}v_i$. The current amplifier model in the same table gives an open-circuit output voltage of $A_{is}i_iR_o$. Equating these two values and noting that $i_i = v_i/R_i$ gives

$$A_{vo} = A_{is}\left(\frac{R_o}{R_i}\right) \tag{1.14}$$

Similarly, we can show that

$$A_{vo} = G_mR_o \tag{1.15}$$

and

$$A_{vo} = \frac{R_m}{R_i} \tag{1.16}$$

The expressions in Eqs. (1.14) to (1.16) can be used to relate any two of the gain parameters A_{vo}, A_{is}, G_m, and R_m.

1.5.5 Determining R_i and R_o

From the amplifier circuit models given in Table 1.1, we observe that the input resistance R_i of the amplifier can be determined by applying an input voltage v_i and measuring (or calculating) the input current i_i; that is, $R_i = v_i/i_i$. The output resistance is found as the ratio of the open-circuit output voltage to the short-circuit output current. Alternatively, the output resistance can be found by eliminating the input signal source (then i_i and v_i will both be zero) and applying a voltage signal v_x to the output of the amplifier, as shown in Fig. 1.18. If we denote the current drawn from v_x *into* the output terminals as i_x (note that i_x is opposite in direction to i_o), then $R_o = v_x/i_x$. Although these techniques are conceptually correct, in actual practice more refined methods are employed in measuring R_i and R_o.

1.5.6 Unilateral Models

The amplifier models considered above are **unilateral**; that is, signal flow is unidirectional, from input to output. Most real amplifiers show some reverse transmission, which is usually undesirable but must nonetheless be modeled. We shall not pursue this point further at this time except to mention that more complete models for linear two-port networks are given in Appendix C. Also, in later chapters, we will find it necessary in certain cases to augment the models of Table 1.1 to take into account the nonunilateral nature of some transistor amplifiers.

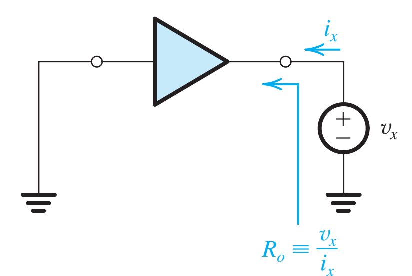


Figure 1.18 Determining the output resistance.

Example 1.4

The **bipolar junction transistor (BJT)**, which will be studied in Chapter 6, is a three-terminal device that when powered up by a dc source (battery) and operated with small signals can be modeled by the linear circuit shown in Fig. 1.19(a). The three terminals are the **base (B)**, the **emitter (E)**, and the **collector (C)**. The heart of the model is a transconductance amplifier represented by an input resistance between B and E (denoted r_π), a short-circuit transconductance g_m, and an output resistance r_o.

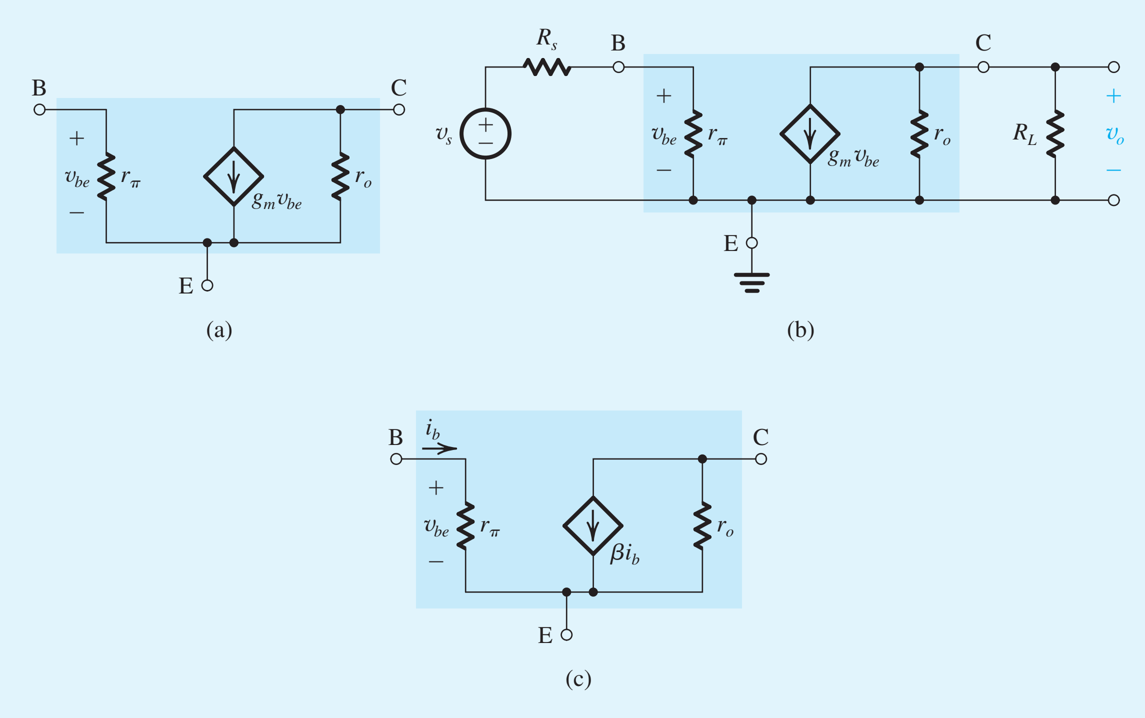


Figure 1.19 **(a)** Small-signal circuit model for a bipolar junction transistor (BJT). **(b)** The BJT connected as an amplifier with the emitter as a common terminal between input and output (called a common-emitter amplifier). **(c)** An alternative small-signal circuit model for the BJT.

(a) With the emitter used as a common terminal between input and output, Fig. 1.19(b) shows a transistor amplifier known as a **common-emitter** or **grounded-emitter** circuit. Derive an expression for the voltage gain v_o/v_s, and evaluate its magnitude for the case $R_s = 5\,\text{k}\Omega$, $r_\pi = 2.5\ \text{k}\Omega$, $g_m = 40\ \text{mA/V}$, $r_o = 100\ \text{k}\Omega$, and $R_L = 5\ \text{k}\Omega$. What would the gain value be if the effect of r_o were neglected?

(b) An alternative model for the transistor in which a current amplifier rather than a transconductance amplifier is utilized is shown in Fig. 1.19(c). What must the short-circuit current gain β be? Give both an expression and a value.

Solution

(a) Refer to Fig. 1.19(b). We use the voltage-divider rule to determine the fraction of input signal that appears at the amplifier input as

$$v_{be} = v_s \frac{r_\pi}{r_\pi + R_s} \tag{1.17}$$

Next we determine the output voltage v_o by multiplying the current $(g_m v_{be})$ by the resistance $(R_L \| r_o)$,

$$v_o = -g_m v_{be}(R_L \| r_o) \tag{1.18}$$

Substituting for v_{be} from Eq. (1.17) yields the voltage-gain expression

$$\frac{v_o}{v_s} = -\frac{r_\pi}{r_\pi + R_s} g_m (R_L \| r_o) \tag{1.19}$$

Observe that the gain is negative, indicating that this amplifier is inverting. For the given component values,

$$\begin{aligned}\frac{v_o}{v_s} &= -\frac{2.5}{2.5+5} \times 40 \times (5 \| 100) \\ &= -63.5 \text{ V/V}\end{aligned}$$

Neglecting the effect of r_o, we obtain

$$\begin{aligned}\frac{v_o}{v_s} &\simeq -\frac{2.5}{2.5+5} \times 40 \times 5 \\ &= -66.7 \text{ V/V}\end{aligned}$$

which is quite close to the value obtained including r_o. This is not surprising, since $r_o \gg R_L$.

(b) For the model in Fig. 1.19(c) to be equivalent to that in Fig. 1.19(a),

$$\beta i_b = g_m v_{be}$$

But $i_b = v_{be}/r_\pi$; thus,

$$\beta = g_m r_\pi$$

For the values given,

$$\begin{aligned}\beta &= 40 \text{ mA/V} \times 2.5 \text{ k}\Omega \\ &= 100 \text{ A/A}\end{aligned}$$

EXERCISES

1.18 Consider a current amplifier having the model shown in the second row of Table 1.1. Let the amplifier be fed with a signal current-source i_s having a resistance R_s, and let the output be connected to a load resistance R_L. Show that the overall current gain is given by

$$\frac{i_o}{i_s} = A_{is}\frac{R_s}{R_s + R_i}\frac{R_o}{R_o + R_L}$$

1.19 Consider the transconductance amplifier whose model is shown in the third row of Table 1.1. Let a voltage signal source v_s with a source resistance R_s be connected to the input and a load resistance R_L be connected to the output. Show that the overall voltage gain is given by

$$\frac{v_o}{v_s} = G_m\frac{R_i}{R_i + R_s}(R_o \| R_L)$$

1.20 Consider a transresistance amplifier having the model shown in the fourth row of Table 1.1. Let the amplifier be fed with a signal current source i_s having a resistance R_s, and let the output be connected to a load resistance R_L. Show that the overall gain is given by

$$\frac{v_o}{i_s} = R_m\frac{R_s}{R_s + R_i}\frac{R_L}{R_L + R_o}$$

1.21 Find the input resistance between terminals B and G in the circuit shown in Fig. E1.21. The voltage v_x is a test voltage with the input resistance R_{in} defined as $R_{\text{in}} \equiv v_x/i_x$.

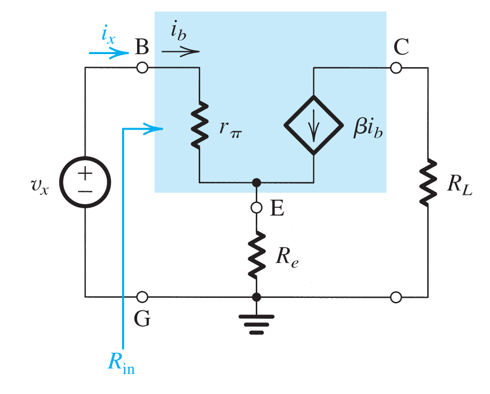


Figure E1.21

Ans. $R_{\text{in}} = r_\pi + (\beta + 1)R_e$

1.6 Frequency Response of Amplifiers[2]

From Section 1.2 we know that the input signal to an amplifier can always be expressed as the sum of sinusoidal signals. It follows that an important characterization of an amplifier is in terms of its response to input sinusoids of different frequencies. Such a characterization of amplifier performance is known as the amplifier frequency response.

1.6.1 Measuring the Amplifier Frequency Response

We shall introduce the subject of amplifier frequency response by showing how it can be measured. Figure 1.20 depicts a linear voltage amplifier fed at its input with a sine-wave signal of amplitude V_i and frequency ω. As the figure indicates, the signal measured at the amplifier output also is sinusoidal with exactly the same frequency ω. This is an important point to note: *Whenever a sine-wave signal is applied to a linear circuit, the resulting output is sinusoidal with the same frequency as the input*. In fact, the sine wave is the only signal that does not change shape as it passes through a linear circuit. Observe, however, that the output sinusoid will in general have a different amplitude and will be shifted in phase relative to the input. The ratio of the amplitude of the output sinusoid (V_o) to the amplitude of the input sinusoid (V_i) is the magnitude of the amplifier gain (or transmission) at the test frequency ω. Also, the angle ϕ is the phase of the amplifier transmission at the test frequency ω. If we denote the **amplifier transmission**, or **transfer function** as it is more commonly known, by $T(\omega)$, then

$$|T(\omega)| = \frac{V_o}{V_i}$$

$$\angle T(\omega) = \phi$$

The response of the amplifier to a sinusoid of frequency ω is completely described by $|T(\omega)|$ and $\angle T(\omega)$. Now, to obtain the complete frequency response of the amplifier we simply change the frequency of the input sinusoid and measure the new value for $|T|$ and $\angle T$. The end result will be a table and/or graph of gain magnitude [$|T(\omega)|$] versus frequency and a table and/or graph of phase angle [$\angle T(\omega)$] versus frequency. These two plots together constitute the frequency response of the amplifier; the first is known as the **magnitude** or **amplitude**

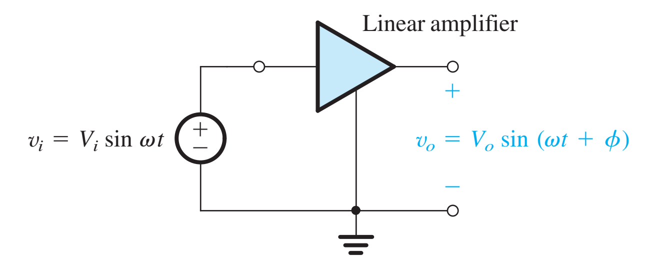


Figure 1.20 Measuring the frequency response of a linear amplifier: At the test frequency, the amplifier gain is characterized by its magnitude (V_o/V_i) and phase ϕ.

[2]Except for its use in the study of the frequency response of op-amp circuits in Sections 2.5 and 2.7, the material in this section will not be needed in a substantial manner until Chapter 10.

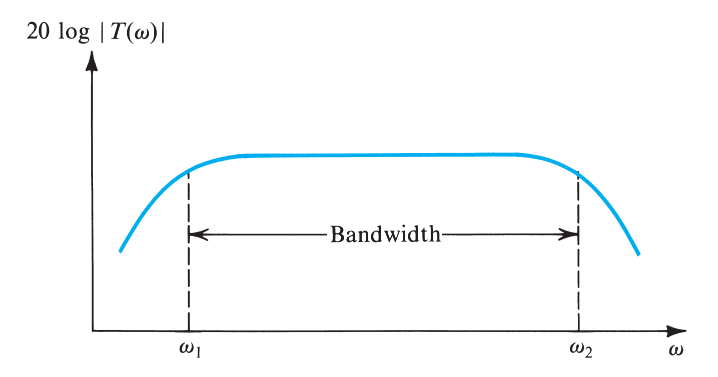


Figure 1.21 Typical magnitude response of an amplifier: $|T(\omega)|$ is the magnitude of the amplifier transfer function—that is, the ratio of the output $V_o(\omega)$ to the input $V_i(\omega)$.

response, and the second is the **phase response**. Finally, we should mention that it is a common practice to express the magnitude of transmission in decibels and thus plot 20 log $|T(\omega)|$ versus frequency.

1.6.2 Amplifier Bandwidth

Figure 1.21 shows the magnitude response of an amplifier. It indicates that the gain is almost constant over a wide frequency range, roughly between ω_1 and ω_2. Signals whose frequencies are below ω_1 or above ω_2 will experience lower gain, with the gain decreasing as we move farther away from ω_1 and ω_2. The band of frequencies over which the gain of the amplifier is almost constant, to within a certain number of decibels (usually 3 dB), is called the **amplifier bandwidth**. Normally the amplifier is designed so that its bandwidth coincides with the spectrum of the signals it is required to amplify. If this were not the case, the amplifier would *distort* the frequency spectrum of the input signal, with different components of the input signal being amplified by different amounts.

1.6.3 Evaluating the Frequency Response of Amplifiers

Above, we described the method used to measure the frequency response of an amplifier. We now briefly discuss the method for analytically obtaining an expression for the frequency response. What we are about to say is just a preview of this important subject, whose detailed study is in Chapter 10.

To evaluate the frequency response of an amplifier, one has to analyze the amplifier equivalent circuit model, taking into account all reactive components.[3] Circuit analysis proceeds in the usual fashion but with inductances and capacitances represented by their reactances. An inductance L has a reactance or impedance $j\omega L$, and a capacitance C has a reactance or impedance $1/j\omega C$ or, equivalently, a susceptance or admittance $j\omega C$. Thus in a *frequency-domain* analysis we deal with impedances and/or admittances. The result of the

[3]Note that in the models considered in previous sections no reactive components were included. These were simplified models and cannot be used alone to predict the amplifier frequency response.

analysis is the amplifier transfer function $T(\omega)$

$$T(\omega) = \frac{V_o(\omega)}{V_i(\omega)}$$

where $V_i(\omega)$ and $V_o(\omega)$ denote the input and output signals, respectively. $T(\omega)$ is generally a complex function whose magnitude $|T(\omega)|$ gives the magnitude of transmission or the magnitude response of the amplifier. The phase of $T(\omega)$ gives the phase response of the amplifier.

In the analysis of a circuit to determine its frequency response, the algebraic manipulations can be considerably simplified by using the **complex frequency variable** s. In terms of s, the impedance of an inductance L is sL and that of a capacitance C is $1/sC$. Replacing the reactive elements with their impedances and performing standard circuit analysis, we obtain the transfer function $T(s)$ as

$$T(s) \equiv \frac{V_o(s)}{V_i(s)}$$

Subsequently, we replace s by $j\omega$ to determine the transfer function for **physical frequencies**, $T(j\omega)$. Note that $T(j\omega)$ is the same function we called $T(\omega)$ above[4]; the additional j is included in order to emphasize that $T(j\omega)$ is obtained from $T(s)$ by replacing s with $j\omega$.

1.6.4 Single-Time-Constant Networks

In analyzing amplifier circuits to determine their frequency response, one is greatly aided by knowledge of the frequency-response characteristics of single-time-constant (STC) networks. An STC network is one that is composed of, or can be reduced to, one reactive component (inductance or capacitance) and one resistance. Examples are shown in Fig. 1.22. An STC network formed of an inductance L and a resistance R has a time constant $\tau = L/R$. The time constant τ of an STC network composed of a capacitance C and a resistance R is given by $\tau = CR$.

Appendix E presents a study of STC networks and their responses to sinusoidal, step, and pulse inputs. Knowledge of this material will be needed at various points throughout this book, and the reader will be encouraged to refer to the appendix. At this point we need in particular the frequency-response results; we will, in fact, briefly discuss this important topic now.

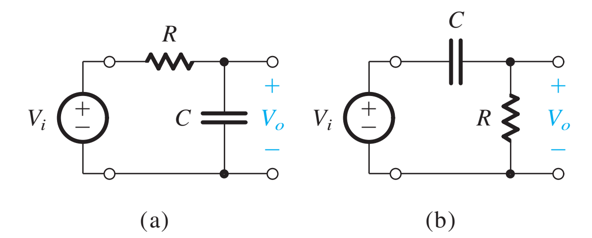


Figure 1.22 Two examples of STC networks: **(a)** a low-pass network and **(b)** a high-pass network.

[4]At this stage, we are using s simply as a shorthand for $j\omega$. We shall not require detailed knowledge of s-plane concepts until Chapter 10. A brief review of s-plane analysis is presented in Appendix F.

Most STC networks can be classified into two categories,[5] **low pass (LP)** and **high pass (HP)**, with each of the two categories displaying distinctly different signal responses. As an example, the STC network shown in Fig. 1.22(a) is of the *low-pass* type and that in Fig. 1.22(b) is of the *high-pass* type. To see the reasoning behind this classification, observe that the transfer function of each of these two circuits can be expressed as a voltage-divider ratio, with the divider composed of a resistor and a capacitor. Now, recalling how the impedance of a capacitor varies with frequency ($Z = 1/j\omega C$), it is easy to see that the transmission of the circuit in Fig. 1.22(a) will decrease with frequency and approach zero as ω approaches ∞. Thus the circuit of Fig. 1.22(a) acts as a **low-pass filter**[6]; it passes low-frequency, sine-wave inputs with little or no attenuation (at $\omega = 0$, the transmission is unity) and attenuates high-frequency input sinusoids. The circuit of Fig. 1.22(b) does the opposite; its transmission is unity at $\omega = \infty$ and decreases as ω is reduced, reaching 0 for $\omega = 0$. The latter circuit, therefore, performs as a **high-pass filter**.

Table 1.2 provides a summary of the frequency-response results for STC networks of both types.[7] Also, sketches of the magnitude and phase responses are given in Figs. 1.23 and 1.24. These frequency-response diagrams are known as **Bode plots**, and the **3-dB frequency** (ω_0) is also known as the **corner frequency**, **break frequency**, or **pole frequency**. The reader is urged to become familiar with this information and to consult Appendix E if further clarifications are needed. In particular, it is important to develop a facility for the rapid

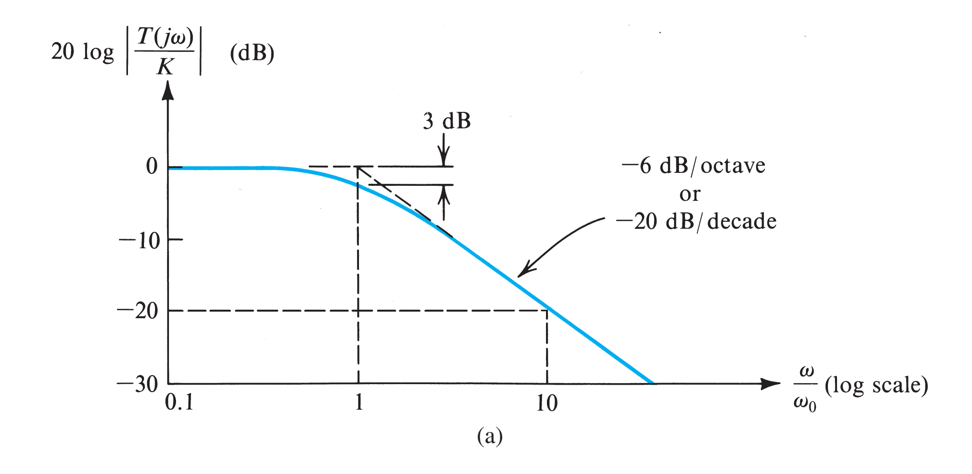
Table 1.2 Frequency Response of STC Networks

	Low-Pass (LP)	High-Pass (HP)
Transfer Function $T(s)$	$\dfrac{K}{1+(s/\omega_0)}$	$\dfrac{Ks}{s+\omega_0}$
Transfer Function (for physical frequencies) $T(j\omega)$	$\dfrac{K}{1+j(\omega/\omega_0)}$	$\dfrac{K}{1-j(\omega_0/\omega)}$
Magnitude Response $\lvert T(j\omega)\rvert$	$\dfrac{\lvert K\rvert}{\sqrt{1+(\omega/\omega_0)^2}}$	$\dfrac{\lvert K\rvert}{\sqrt{1+(\omega_0/\omega)^2}}$
Phase Response $\angle T(j\omega)$	$-\tan^{-1}(\omega/\omega_0)$	$\tan^{-1}(\omega_0/\omega)$
Transmission at $\omega = 0$ (dc)	K	0
Transmission at $\omega = \infty$	0	K
3-dB Frequency	$\omega_0 = 1/\tau$; $\tau \equiv$ time constant $\tau = CR$ or L/R	
Bode Plots	in Fig. 1.23	in Fig. 1.24

[5] An important exception is the **all-pass** STC network studied in Chapter 17.

[6] A filter is a circuit that passes signals in a specified frequency band (the filter passband) and stops or severely attenuates (filters out) signals in another frequency band (the filter stopband). Filters will be studied in Chapter 17.

[7] The transfer functions in Table 1.2 are given in general form. For the circuits of Fig. 1.22, $K = 1$ and $\omega_0 = 1/CR$.



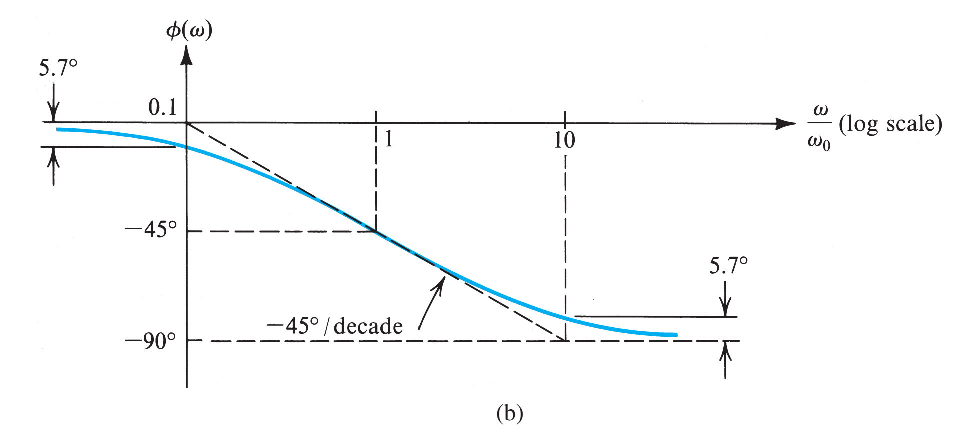
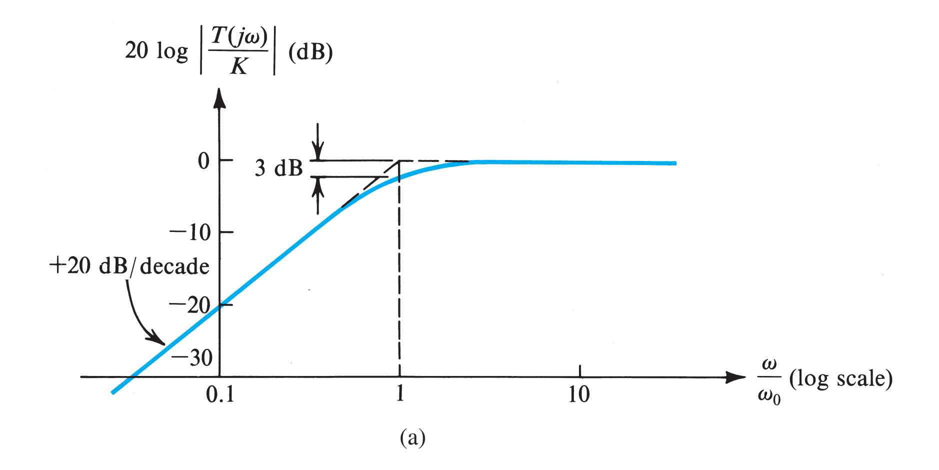


Figure 1.23 **(a)** Magnitude and **(b)** phase response of STC networks of the low-pass type.

determination of the time constant τ of an STC circuit. The process is very simple: Set the independent voltage or current source to zero; "grab hold" of the two terminals of the reactive element (capacitor C or inductor L); and determine the equivalent resistance R that appears between these two terminals. The time constant is then CR or L/R.

BODE PLOTS:

In the 1930s, while working at Bell Labs, Hendrik Bode devised a simple but accurate method for using linearized asymptotic responses to graph gain and phase shift against frequency on a logarithmic scale. Such gain and phase presentations, together called Bode plots, have enormous importance in the design and analysis of the frequency-dependent behavior of systems large and small.



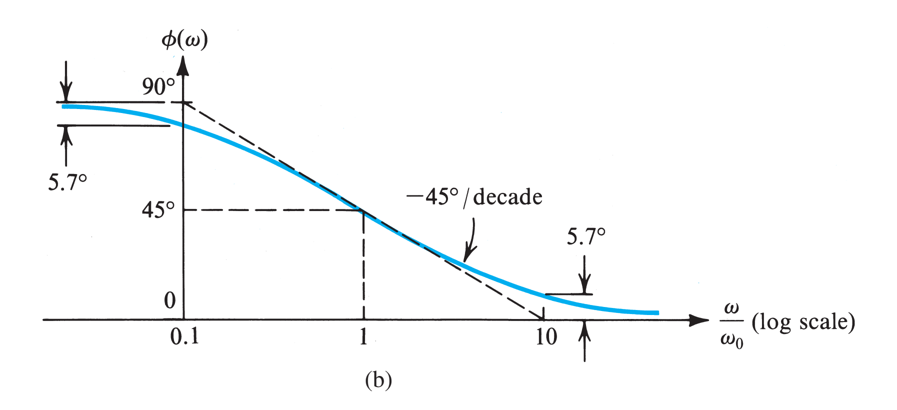


Figure 1.24 **(a)** Magnitude and **(b)** phase response of STC networks of the high-pass type.

Example 1.5

Figure 1.25 shows a voltage amplifier having an input resistance R_i, an input capacitance C_i, a gain factor μ, and an output resistance R_o. The amplifier is fed with a voltage source V_s having a source resistance R_s, and a load of resistance R_L is connected to the output.

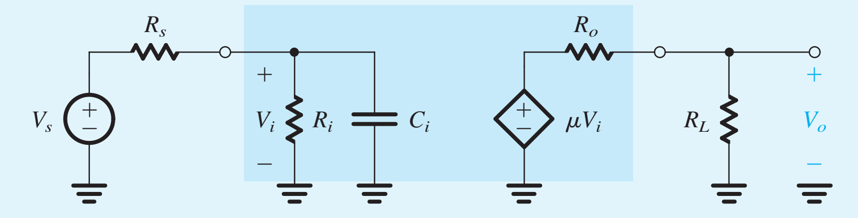


Figure 1.25 Circuit for Example 1.5.

(a) Derive an expression for the amplifier voltage gain V_o/V_s as a function of frequency. From this find expressions for the dc gain and the 3-dB frequency.
(b) Calculate the values of the dc gain, the 3-dB frequency, and the frequency at which the gain becomes 0 dB (i.e., unity) for the case $R_s = 20\text{ k}\Omega$, $R_i = 100\text{ k}\Omega$, $C_i = 60\text{ pF}$, $\mu = 144\text{ V/V}$, $R_o = 200\ \Omega$, and $R_L = 1\text{ k}\Omega$.
(c) Find $v_o(t)$ for each of the following inputs:
   (i) $v_i = 0.1 \sin 10^2 t$, V
   (ii) $v_i = 0.1 \sin 10^5 t$, V
   (iii) $v_i = 0.1 \sin 10^6 t$, V
   (iv) $v_i = 0.1 \sin 10^8 t$, V

Solution

(a) Utilizing the voltage-divider rule, we can express V_i in terms of V_s as follows

$$V_i = V_s \frac{Z_i}{Z_i + R_s}$$

where Z_i is the amplifier input impedance. Since Z_i is composed of two parallel elements, it is obviously easier to work in terms of $Y_i = 1/Z_i$. Toward that end we divide the numerator and denominator by Z_i, thus obtaining

$$\begin{aligned} V_i &= V_s \frac{1}{1 + R_s Y_i} \\ &= V_s \frac{1}{1 + R_s[(1/R_i) + sC_i]} \end{aligned}$$

Thus,

$$\frac{V_i}{V_s} = \frac{1}{1 + (R_s/R_i) + sC_i R_s}$$

This expression can be put in the standard form for a low-pass STC network (see the top line of Table 1.2) by extracting $[1 + (R_s/R_i)]$ from the denominator; thus we have

$$\frac{V_i}{V_s} = \frac{1}{1 + (R_s/R_i)} \frac{1}{1 + sC_i[(R_s R_i)/(R_s + R_i)]} \tag{1.20}$$

At the output side of the amplifier we can use the voltage-divider rule to write

$$V_o = \mu V_i \frac{R_L}{R_L + R_o}$$

This equation can be combined with Eq. (1.20) to obtain the amplifier transfer function as

$$\frac{V_o}{V_s} = \mu \frac{1}{1 + (R_s/R_i)} \frac{1}{1 + (R_o/R_L)} \frac{1}{1 + sC_i[(R_s R_i)/(R_s + R_i)]} \tag{1.21}$$

Example 1.5 *continued*

We note that only the last factor in this expression is new (compared with the expression derived in the last section). This factor is a result of the input capacitance C_i, with the time constant being

$$\tau = C_i \frac{R_s R_i}{R_s + R_i} = C_i (R_s \| R_i) \tag{1.22}$$

We could have obtained this result by inspection: From Fig. 1.25 we see that the input circuit is an STC network and that its time constant can be found by reducing V_s to zero, with the result that the resistance seen by C_i is R_i in parallel with R_s. The transfer function in Eq. (1.21) is of the form $K/(1+(s/\omega_o))$, which corresponds to a low-pass STC network. The dc gain is found as

$$K \equiv \frac{V_o}{V_s}(s=0) = \mu \frac{1}{1+(R_s/R_i)} \frac{1}{1+(R_o/R_L)} \tag{1.23}$$

The 3-dB frequency ω_0 can be found from

$$\omega_0 = \frac{1}{\tau} = \frac{1}{C_i(R_s \| R_i)} \tag{1.24}$$

Since the frequency response of this amplifier is of the low-pass STC type, the Bode plots for the gain magnitude and phase will take the form shown in Fig. 1.23, where K is given by Eq. (1.23) and ω_0 is given by Eq. (1.24).

(b) Substituting the numerical values given into Eq. (1.23) results in

$$K = 144 \frac{1}{1+(20/100)} \frac{1}{1+(200/1000)} = 100 \text{ V/V}$$

Thus the amplifier has a dc gain of 40 dB. Substituting the numerical values into Eq. (1.24) gives the 3-dB frequency

$$\omega_0 = \frac{1}{60 \text{ pF} \times (20 \text{ k}\Omega // 100 \text{ k}\Omega)}$$

$$= \frac{1}{60 \times 10^{-12} \times (20 \times 100/(20+100)) \times 10^3} = 10^6 \text{ rad/s}$$

Thus,

$$f_0 = \frac{10^6}{2\pi} = 159.2 \text{ kHz}$$

Since the gain falls off at the rate of –20 dB/decade, starting at ω_0 (see Fig. 1.23a) the gain will reach 0 dB in two decades (a factor of 100); thus we have

$$\text{Unity-gain frequency} = 100 \times \omega_0 = 10^8 \text{ rad/s or } 15.92 \text{ MHz}$$

(c) To find $v_o(t)$ we need to determine the gain magnitude and phase at 10^2, 10^5, 10^6, and 10^8 rad/s. This can be done either approximately utilizing the Bode plots of Fig. 1.23 or exactly utilizing the expression for the amplifier transfer function,

$$T(j\omega) \equiv \frac{V_o}{V_s}(j\omega) = \frac{100}{1 + j(\omega/10^6)}$$

We shall do both:

(i) For $\omega = 10^2$ rad/s, which is $(\omega_0/10^4)$, the Bode plots of Fig. 1.23 suggest that $|T| = K = 100$ and $\phi = 0°$. The transfer function expression gives $|T| \simeq 100$ and $\phi = -\tan^{-1} 10^{-4} \simeq 0°$. Thus,

$$v_o(t) = 10 \sin 10^2 t, \ \text{V}$$

(ii) For $\omega = 10^5$ rad/s, which is $(\omega_0/10)$, the Bode plots of Fig. 1.23 suggest that $|T| \simeq K = 100$ and $\phi = -5.7°$. The transfer function expression gives $|T| = 99.5$ and $\phi = -\tan^{-1} 0.1 = -5.7°$. Thus,

$$v_o(t) = 9.95 \sin(10^5 t - 5.7°), \ \text{V}$$

(iii) For $\omega = 10^6$ rad/s $= \omega_0$, $|T| = 100/\sqrt{2} = 70.7$ V/V or 37 dB and $\phi = -45°$. Thus,

$$v_o(t) = 7.07 \sin(10^6 t - 45°), \ \text{V}$$

(iv) For $\omega = 10^8$ rad/s, which is $(100\,\omega_0)$, the Bode plots suggest that $|T| = 1$ and $\phi = -90°$. The transfer function expression gives $|T| \simeq 1$ and $\phi = -\tan^{-1} 100 = -89.4°$. Thus,

$$v_o(t) = 0.1 \sin(10^8 t - 89.4°), \ \text{V}$$

1.6.5 Classification of Amplifiers Based on Frequency Response

Amplifiers can be classified based on the shape of their magnitude-response curve. Figure 1.26 shows typical frequency-response curves for various amplifier types. In Fig. 1.26(a) the gain remains constant over a wide frequency range, but falls off at low and high frequencies. This type of frequency response is common in audio amplifiers.

As will be shown in later chapters, **internal capacitances** in the device (a transistor) cause the falloff of gain at high frequencies, just as C_i did in the circuit of Example 1.5. On the other hand, the falloff of gain at low frequencies is usually caused by **coupling capacitors** used to connect one amplifier stage to another, as indicated in Fig. 1.27. This practice is usually adopted to simplify the design process of the different stages. The coupling capacitors are usually chosen quite large (a fraction of a microfarad to a few tens of microfarads) so that their reactance (impedance) is small at the frequencies of interest. Nevertheless, at sufficiently low frequencies the reactance of a coupling capacitor will become large enough to cause part of the signal being coupled to appear as a voltage drop across the coupling capacitor, thus not reaching the subsequent stage. Coupling capacitors will thus cause loss of gain at low

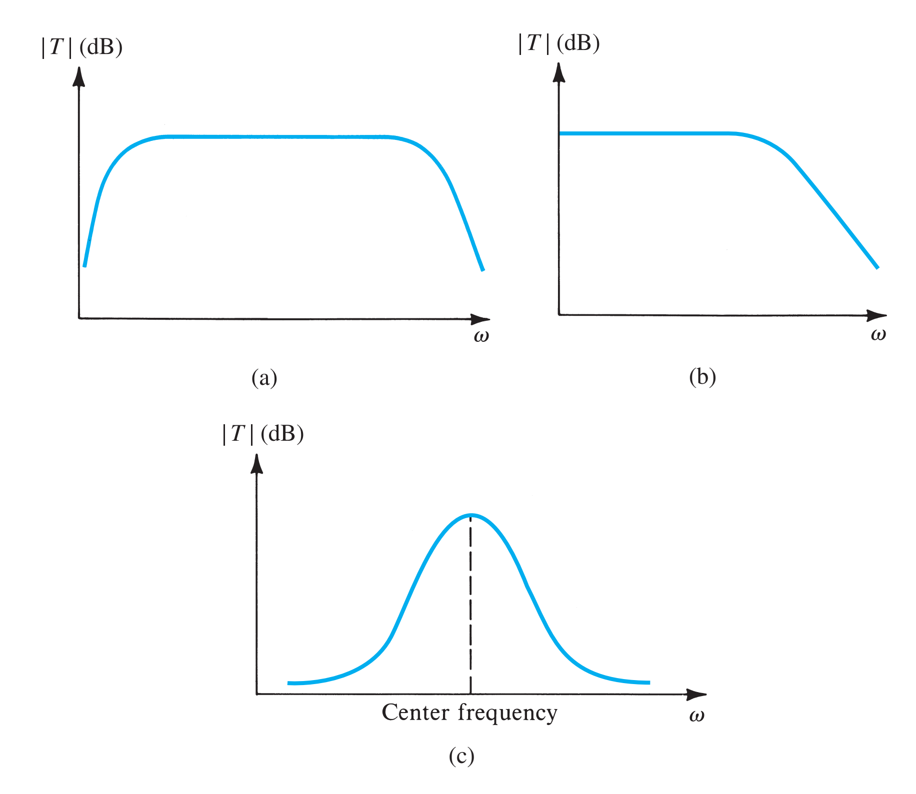


Figure 1.26 Frequency response for **(a)** a capacitively coupled amplifier, **(b)** a direct-coupled amplifier, and **(c)** a tuned or bandpass amplifier.

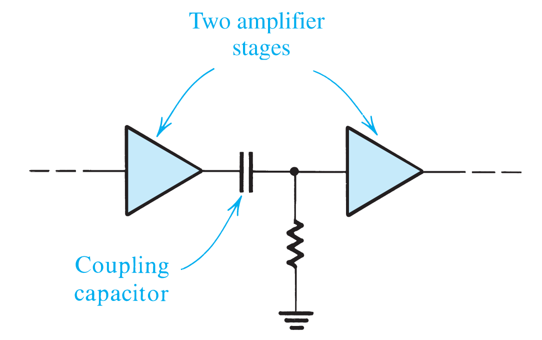


Figure 1.27 Use of a capacitor to couple amplifier stages.

frequencies and cause the gain to be zero at dc. This is not at all surprising, since from Fig. 1.27 we observe that the coupling capacitor, acting together with the input resistance of the subsequent stage, forms a high-pass STC circuit. It is the frequency response of this high-pass circuit that accounts for the shape of the amplifier frequency response in Fig. 1.26(a) at the low-frequency end.

There are many applications in which it is important that the amplifier maintain its gain at low frequencies down to dc. Furthermore, monolithic integrated-circuit (IC) technology does not allow the fabrication of large coupling capacitors. Thus IC amplifiers are usually designed as **directly coupled** or **dc amplifiers** (as opposed to **capacitively coupled**, or **ac amplifiers**).

Figure 1.26(b) shows the frequency response of a dc amplifier. Such a frequency response characterizes what is referred to as a **low-pass amplifier**.

In a number of applications, such as in the design of radio and TV receivers, the need arises for an amplifier whose frequency response peaks around a certain frequency (called the **center frequency**) and falls off on both sides of this frequency, as shown in Fig. 1.26(c). Amplifiers with such a response are called **tuned amplifiers**, **bandpass amplifiers**, or **bandpass filters**. A tuned amplifier forms the heart of the front-end or tuner of a communication receiver; by adjusting its center frequency to coincide with the frequency of a desired communications channel (e.g., a radio station), the signal of this particular channel can be received while those of other channels are attenuated or filtered out.

EXERCISES

1.22 Consider a voltage amplifier having a frequency response of the low-pass STC type with a dc gain of 60 dB and a 3-dB frequency of 1000 Hz. Find the gain in dB at $f = 10$ Hz, 10 kHz, 100 kHz, and 1 MHz.

Ans. 60 dB; 40 dB; 20 dB; 0 dB

D1.23 Consider a transconductance amplifier having the model shown in Table 1.1 with $R_i = 5$ kΩ, $R_o = 50$ kΩ, and $G_m = 10$ mA/V. If the amplifier load consists of a resistance R_L in parallel with a capacitance C_L, convince yourself that the voltage transfer function realized, V_o/V_i, is of the low-pass STC type. What is the lowest value that R_L can have while a dc gain of at least 40 dB is obtained? With this value of R_L connected, find the highest value that C_L can have while a 3-dB bandwidth of at least 100 kHz is obtained.

Ans. 12.5 kΩ; 159.2 pF

D1.24 Consider the situation illustrated in Fig. 1.27. Let the output resistance of the first voltage amplifier be 1 kΩ and the input resistance of the second voltage amplifier (including the resistor shown) be 9 kΩ. The resulting equivalent circuit is shown in Fig. E1.24. Convince yourself that V_2/V_s is a high-pass STC function. What is the smallest value for C that will ensure that the 3-dB frequency is not higher than 100 Hz?

Ans. 0.16 μF

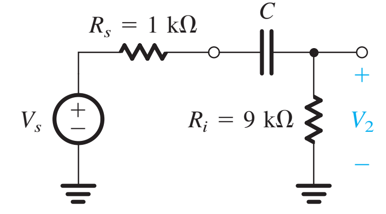


Figure E1.24

Summary

- An electrical signal source can be represented in either the Thévenin form (a voltage source v_s in series with a source resistance R_s) or the Norton form (a current source i_s in parallel with a source resistance R_s). The Thévenin voltage v_s is the open-circuit voltage between the source terminals; the Norton current i_s is equal to the short-circuit current between the source terminals. For the two representations to be equivalent, v_s and $R_s i_s$ must be equal.
- A signal can be represented either by its waveform versus time or as the sum of sinusoids. The latter representation is known as the frequency spectrum of the signal.
- The sine-wave signal is completely characterized by its peak value (or rms value, which is the peak/$\sqrt{2}$), its frequency (ω in rad/s or f in Hz; $\omega = 2\pi f$ and $f = 1/T$, where T is the period in seconds), and its phase with respect to an arbitrary reference time.
- Analog signals have magnitudes that can assume any value. Electronic circuits that process analog signals are called analog circuits. Sampling the magnitude of an analog signal at discrete instants of time and representing each signal sample by a number results in a digital signal. Digital signals are processed by digital circuits.
- The simplest digital signals are obtained when the binary system is used. An individual digital signal then assumes one of only two possible values: low and high (say, 0 V and +5 V), corresponding to logic 0 and logic 1, respectively.
- An analog-to-digital converter (ADC) provides at its output the digits of the binary number representing the analog signal sample applied to its input. The output digital signal can then be processed using digital circuits. Refer to Fig. 1.10 and Eq. (1.3).
- The transfer characteristic, v_O versus v_I, of a linear amplifier is a straight line with a slope equal to the voltage gain. Refer to Fig. 1.12.
- Amplifiers increase the signal power and thus require dc power supplies for their operation.
- The amplifier voltage gain can be expressed as a ratio A_v in V/V or in decibels, $20 \log |A_v|$, dB. Similarly, for current gain: A_i A/A or $20 \log |A_i|$, dB. For power gain: A_p W/W or $10 \log A_p$, dB.
- Depending on the signal to be amplified (voltage or current) and on the desired form of output signal (voltage or current), there are four basic amplifier types: voltage, current, transconductance, and transresistance amplifiers. For the circuit models and ideal characteristics of these four amplifier types, refer to Table 1.1. A given amplifier can be modeled by any one of the four models, in which case their parameters are related by the formulas in Eqs. (1.14) to (1.16).
- A sinusoid is the only signal whose waveform is unchanged through a linear circuit. Sinusoidal signals are used to measure the frequency response of amplifiers.
- The transfer function $T(s) \equiv V_o(s)/V_i(s)$ of a voltage amplifier can be determined from circuit analysis. Substituting $s = j\omega$ gives $T(j\omega)$, whose magnitude $|T(j\omega)|$ is the magnitude response, and whose phase $\phi(\omega)$ is the phase response, of the amplifier.
- Amplifiers are classified according to the shape of their frequency response, $|T(j\omega)|$. Refer to Fig. 1.26.
- Single-time-constant (STC) networks are those networks that are composed of, or can be reduced to, one reactive component (L or C) and one resistance (R). The time constant τ is either L/R or CR.
- STC networks can be classified into two categories: low pass (LP) and high pass (HP). LP networks pass dc and low frequencies and attenuate high frequencies. The opposite is true for HP networks.
- The gain of an LP (HP) STC circuit drops by 3 dB below the zero-frequency (infinite-frequency) value at a frequency $\omega_0 = 1/\tau$. At high frequencies (low frequencies) the gain falls off at the rate of 6 dB/octave or 20 dB/decade. Refer to Table 1.2 on page 36 and Figs. 1.23 and 1.24. Further details are given in Appendix E.

Circuit Basics

As a review of the basics of circuit analysis and in order for the readers to gauge their preparedness for the study of electronic circuits, this section presents a number of relevant circuit analysis problems. For a summary of Thévenin's and Norton's theorems, refer to Appendix D. The problems are grouped in appropriate categories.

Resistors and Ohm's Law

1.1 Ohm's law relates V, I, and R for a resistor. For each of the situations following, find the missing item:

(a) $R = 1\ \text{k}\Omega, V = 5\ \text{V}$
(b) $V = 5\ \text{V}, I = 1\ \text{mA}$
(c) $R = 10\ \text{k}\Omega,\ I = 0.1\ \text{mA}$
(d) $R = 100\ \Omega, V = 1\ \text{V}$

Note: Volts, milliamps, and kilohms constitute a consistent set of units.

1.2 Measurements taken on various resistors are shown below. For each, calculate the power dissipated in the resistor and the power rating necessary for safe operation using standard components with power ratings of 1/8 W, 1/4 W, 1/2 W, 1 W, or 2 W:

(a) 1 kΩ conducting 20 mA
(b) 1 kΩ conducting 40 mA
(c) 100 kΩ conducting 1 mA
(d) 10 kΩ conducting 4 mA
(e) 1 kΩ dropping 20 V
(f) 1 kΩ dropping 11 V

1.3 Ohm's law and the power law for a resistor relate V, I, R, and P, making only two variables independent. For each pair identified below, find the other two:

(a) $R = 1\ \text{k}\Omega,\ I = 5\ \text{mA}$
(b) $V = 5\ \text{V}, I = 1\ \text{mA}$
(c) $V = 10\ \text{V}, P = 100\ \text{mW}$
(d) $I = 0.1\ \text{mA}, P = 1\ \text{mW}$
(e) $R = 1\ \text{k}\Omega,\ P = 1\ \text{W}$

Combining Resistors

1.4 You are given three resistors whose values are 10 kΩ, 20 kΩ, and 40 kΩ. How many different resistances can you create using series and parallel combinations of these three? List them in value order, lowest first. Be thorough and organized. (*Hint*: In your search, first consider all parallel combinations, then consider series combinations, and then consider series-parallel combinations, of which there are two kinds.)

1.5 In the analysis and test of electronic circuits, it is often useful to connect one resistor in parallel with another to obtain a nonstandard value, one which is smaller than the smaller of the two resistors. Often, particularly during circuit testing, one resistor is already installed, in which case the second, when connected in parallel, is said to "shunt" the first. If the original resistor is 10 kΩ, what is the value of the shunting resistor needed to reduce the combined value by 1%, 5%, 10%, and 50%? What is the result of shunting a 10-kΩ resistor by 1 MΩ? By 100 kΩ? By 10 kΩ?

Voltage Dividers

1.6 Figure P1.6(a) shows a two-resistor voltage divider. Its function is to generate a voltage V_O (smaller than the power-supply voltage V_{DD}) at its output node X. The circuit looking back at node X is equivalent to that shown in Fig. P1.6(b). Observe that this is the Thévenin equivalent of the voltage-divider circuit. Find expressions for V_O and R_O.

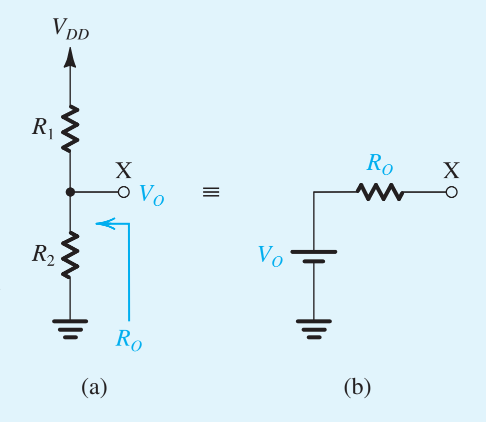


Figure P1.6

1.7 A two-resistor voltage divider employing a 2-kΩ and a 3-kΩ resistor is connected to a 5-V ground-referenced power supply to provide a 2-V voltage. Sketch the circuit. Assuming exact-valued resistors, what output voltage (measured to ground) and equivalent output resistance result? If the resistors used are not ideal but have a ±5% manufacturing tolerance, what are the extreme output voltages and resistances that can result?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

D 1.8 You are given three resistors, each of 10 kΩ, and a 9-V battery whose negative terminal is connected to ground. With a voltage divider using some or all of your resistors, how many positive-voltage sources of magnitude less than 9 V can you design? List them in order, smallest first. What is the output resistance (i.e., the Thévenin resistance) of each?

D *1.9 Two resistors, with nominal values of 4.7 kΩ and 10 kΩ, are used in a voltage divider with a +15-V supply to create a nominal +5-V output. Assuming the resistor values to be exact, what is the actual output voltage produced? Which resistor must be shunted (paralleled) by what third resistor to create a voltage-divider output of 5.00 V? If an output resistance of exactly 3.33 kΩ is also required, what do you suggest?

Current Dividers

1.10 Current dividers play an important role in circuit design. Therefore it is important to develop a facility for dealing with current dividers in circuit analysis. Figure P1.10 shows a two-resistor current divider fed with an ideal current source I. Show that

$$I_1 = \frac{R_2}{R_1 + R_2} I$$

$$I_2 = \frac{R_1}{R_1 + R_2} I$$

and find the voltage V that develops across the current divider.

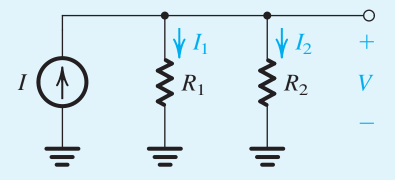


Figure P1.10

D 1.11 Design a simple current divider that will reduce the current provided to a 10-kΩ load to one-third of that available from the source.

D 1.12 A designer searches for a simple circuit to provide one-fifth of a signal current I to a load resistance R. Suggest a solution using one resistor. What must its value be? What is the input resistance of the resulting current divider? For a particular value R, the designer discovers that the otherwise-best-available resistor is 10% too high. Suggest two circuit topologies using one additional resistor that will solve this problem. What is the value of the resistor required in each case? What is the input resistance of the current divider in each case?

D 1.13 A particular electronic signal source generates currents in the range 0 mA to 0.5 mA under the condition that its load voltage not exceed 1 V. For loads causing more than 1 V to appear across the generator, the output current is no longer assured but will be reduced by some unknown amount. This circuit limitation, occurring, for example, at the peak of a sine-wave signal, will lead to undesirable signal distortion that must be avoided. If a 10-kΩ load is to be connected, what must be done? What is the name of the circuit you must use? How many resistors are needed? What is (are) the(ir) value(s)? What is the range of current through the load?

Thévenin Equivalent Circuits

1.14 For the circuit in Fig. P1.14, find the Thévenin equivalent circuit between terminals (a) 1 and 2, (b) 2 and 3, and (c) 1 and 3.

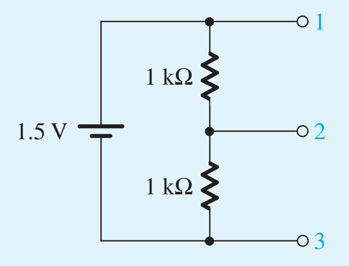


Figure P1.14

1.15 Through repeated application of Thévenin's theorem, find the Thévenin equivalent of the circuit in Fig. P1.15 between node 4 and ground, and hence find the current that flows through a load resistance of 3 kΩ connected between node 4 and ground.

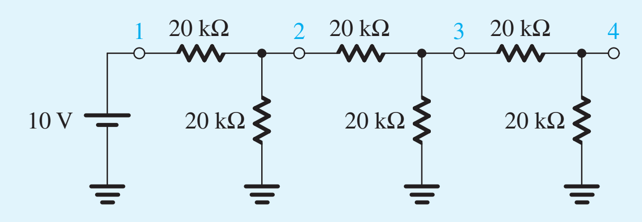


Figure P1.15

Circuit Analysis

1.16 For the circuit shown in Fig. P1.16, find the current in each of the three resistors and the voltage (with respect to ground) at their common node using two methods:

(a) Loop Equations: Define branch currents I_1 and I_2 in R_1 and R_2, respectively; write two equations; and solve them.
(b) Node Equation: Define the node voltage V at the common node; write a single equation; and solve it.

Which method do you prefer? Why?

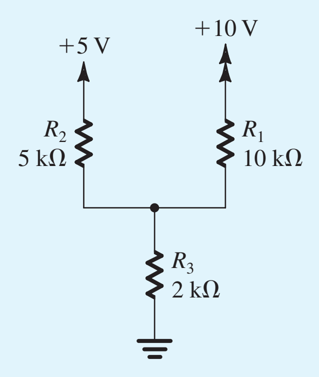


Figure P1.16

1.17 The circuit shown in Fig. P1.17 represents the equivalent circuit of an unbalanced bridge. It is required to calculate the current in the detector branch (R_5) and the voltage across it. Although this can be done by using loop and node equations, a much easier approach is possible: Find the Thévenin equivalent of the circuit to the left of node 1 and the Thévenin equivalent of the circuit to the right of node 2. Then solve the resulting simplified circuit.

***1.18** For the circuit in Fig. P1.18, find the equivalent resistance to ground, R_{eq}. To do this, apply a voltage V_x between terminal X and ground and find the current drawn from V_x. Note that you can use particular special properties of the circuit to get the result directly! Now, if R_4 is raised to 1.2 kΩ, what does R_{eq} become?

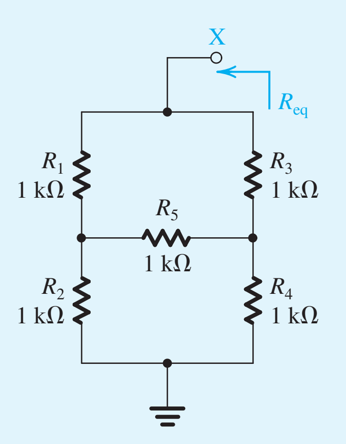


Figure P1.18

1.19 Derive an expression for v_o/v_s for the circuit shown in Fig. P1.19.

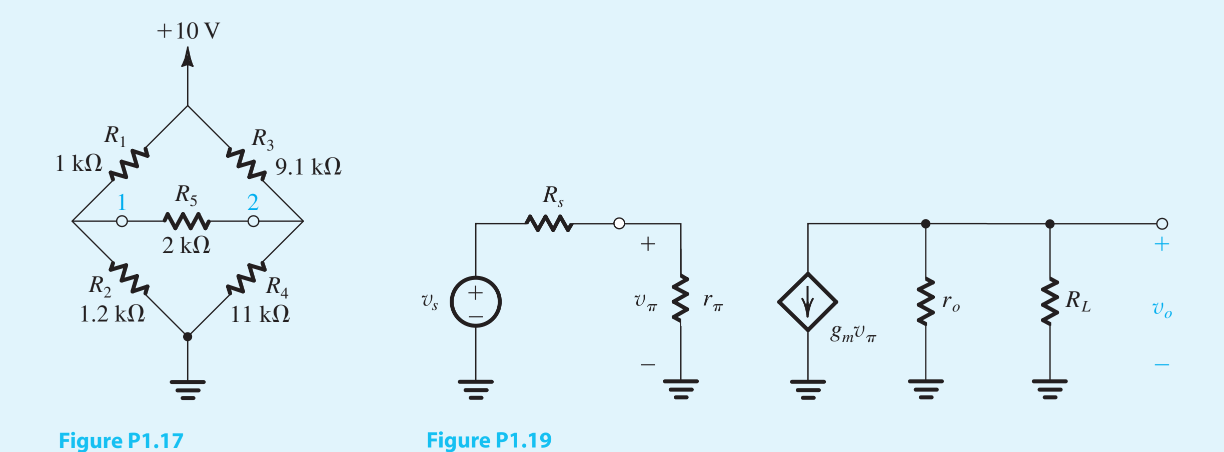


Figure P1.17

Figure P1.19

AC Circuits

1.20 The periodicity of recurrent waveforms, such as sine waves or square waves, can be completely specified using only one of three possible parameters: radian frequency, ω, in radians per second (rad/s); (conventional) frequency, f, in hertz (Hz); or period T, in seconds (s). As well, each of the parameters can be specified numerically in one of several ways: using letter prefixes associated with the basic units, using scientific notation, or using some combination of both. Thus, for example, a particular period may be specified as 100 ns, 0.1 μs, 10^{-1} μs, 10^{5} ps, or 1×10^{-7} s. (For the definition of the various prefixes used in electronics, see Appendix J.) For each of the measures listed below, express the trio of terms in scientific notation associated with the basic unit (e.g., 10^{-7} s rather than 10^{-1} μs).

(a) $T = 10^{-4}$ ms
(b) $f = 1$ GHz
(c) $\omega = 6.28 \times 10^{2}$ rad/s
(d) $T = 10$ s
(e) $f = 60$ Hz
(f) $\omega = 1$ krad/s
(g) $f = 1900$ MHz

1.21 Find the complex impedance, Z, of each of the following basic circuit elements at 60 Hz, 100 kHz, and 1 GHz:

(a) $R = 1$ kΩ
(b) $C = 10$ nF
(c) $C = 10$ pF
(d) $L = 10$ mH
(e) $L = 1$ μH

1.22 Find the complex impedance at 10 kHz of the following networks:

(a) 1 kΩ in series with 10 nF
(b) 10 kΩ in parallel with 0.01 μF
(c) 100 kΩ in parallel with 100 pF
(d) 100 Ω in series with 10 mH

Section 1.1: Signals

1.23 Any given signal source provides an open-circuit voltage, v_{oc}, and a short-circuit current, i_{sc}. For the following sources, calculate the internal resistance, R_s; the Norton current, i_s; and the Thévenin voltage, v_s:

(a) $v_{oc} = 1$ V, $i_{sc} = 0.1$ mA
(b) $v_{oc} = 0.1$ V, $i_{sc} = 1$ μA

1.24 A particular signal source produces an output of 40 mV when loaded by a 100-kΩ resistor and 10 mV when loaded by a 10-kΩ resistor. Calculate the Thévenin voltage, Norton current, and source resistance.

1.25 A temperature sensor is specified to provide 2 mV/°C. When connected to a load resistance of 5 kΩ, the output voltage was measured to change by 10 mV, corresponding to a change in temperature of 10°C. What is the source resistance of the sensor?

1.26 Refer to the Thévenin and Norton representations of the signal source (Fig. 1.1). If the current supplied by the source is denoted i_o and the voltage appearing between the source output terminals is denoted v_o, sketch and clearly label v_o versus i_o for $0 \le i_o \le i_s$.

1.27 The connection of a signal source to an associated signal processor or amplifier generally involves some degree of signal loss as measured at the processor or amplifier input. Considering the two signal-source representations shown in Fig. 1.1, provide two sketches showing each signal-source representation connected to the input terminals (and corresponding input resistance) of a signal processor. What signal-processor input resistance will result in 95% of the open-circuit voltage being delivered to the processor? What input resistance will result in 95% of the short-circuit signal current entering the processor?

Section 1.2: Frequency Spectrum of Signals

1.28 To familiarize yourself with typical values of angular frequency ω, conventional frequency f, and period T, complete the entries in the following table:

Case	ω (rad/s)	f (Hz)	T (s)
a		5×10^{9}	
b	2×10^{9}		
c			1×10^{-10}
d		60	
e	6.28×10^{4}		
f			1×10^{-5}

1.29 For the following peak or rms values of some important sine waves, calculate the corresponding other value:

(a) 117 V rms, a household-power voltage in North America
(b) 33.9 V peak, a somewhat common peak voltage in rectifier circuits
(c) 220 V rms, a household-power voltage in parts of Europe
(d) 220 kV rms, a high-voltage transmission-line voltage in North America

1.30 Give expressions for the sine-wave voltage signals having:

(a) 10-V peak amplitude and 1-kHz frequency
(b) 120-V rms and 60-Hz frequency
(c) 0.2-V peak-to-peak and 2000-rad/s frequency
(d) 100-mV peak and 1-ms period

1.31 Using the information provided by Eq. (1.2) in association with Fig. 1.5, characterize the signal represented by $v(t) = 1/2 + 2/\pi(\sin 2000\pi t + \frac{1}{3}\sin 6000\pi t + \frac{1}{5}\sin 10{,}000\pi t + \cdots)$. Sketch the waveform. What is its average value? Its peak-to-peak value? Its lowest value? Its highest value? Its frequency? Its period?

1.32 Measurements taken of a square-wave signal using a frequency-selective voltmeter (called a spectrum analyzer) show its spectrum to contain adjacent components (spectral lines) at 98 kHz and 126 kHz of amplitudes 63 mV and 49 mV, respectively. For this signal, what would direct measurement of the fundamental show its frequency and amplitude to be? What is the rms value of the fundamental? What are the peak-to-peak amplitude and period of the originating square wave?

1.33 Find the amplitude of a symmetrical square wave of period T that provides the same power as a sine wave of peak amplitude $\hat{V}$ and the same frequency. Does this result depend on equality of the frequencies of the two waveforms?

Section 1.3: Analog and Digital Signals

1.34 Give the binary representation of the following decimal numbers: 0, 6, 11, 28, and 59.

1.35 Consider a 4-bit digital word $b_3b_2b_1b_0$ in a format called signed-magnitude, in which the most significant bit, b_3, is interpreted as a sign bit—0 for positive and 1 for negative values. List the values that can be represented by this scheme. What is peculiar about the representation of zero? For a particular analog-to-digital converter (ADC), each change in b_0 corresponds to a 0.5-V change in the analog input. What is the full range of the analog signal that can be represented? What signed-magnitude digital code results for an input of +2.5 V? For −3.0 V? For +2.7 V? For −2.8 V?

1.36 Consider an N-bit ADC whose analog input varies between 0 and V_{FS} (where the subscript FS denotes "full scale").

(a) Show that the least significant bit (LSB) corresponds to a change in the analog signal of $V_{FS}/(2^N - 1)$. This is the resolution of the converter.
(b) Convince yourself that the maximum error in the conversion (called the quantization error) is half the resolution; that is, the quantization error $= V_{FS}/2(2^N - 1)$.
(c) For $V_{FS} = 5$ V, how many bits are required to obtain a resolution of 2 mV or better? What is the actual resolution obtained? What is the resulting quantization error?

1.37 Figure P1.37 shows the circuit of an N-bit digital-to-analog converter (DAC). Each of the N bits of the digital word to be converted controls one of the switches.

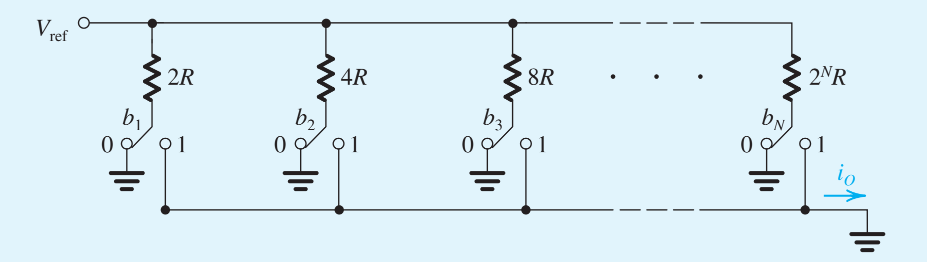


Figure P1.37

When the bit is 0, the switch is in the position labeled 0; when the bit is 1, the switch is in the position labeled 1. The analog output is the current i_O. V_{ref} is a constant reference voltage.

(a) Show that

$$i_O = \frac{V_{ref}}{R}\left(\frac{b_1}{2^1} + \frac{b_2}{2^2} + \cdots + \frac{b_N}{2^N}\right)$$

(b) Which bit is the LSB? Which is the MSB?

(c) For $V_{ref} = 10$ V, $R = 10$ kΩ, and $N = 8$, find the maximum value of i_O obtained. What is the change in i_O resulting from the LSB changing from 0 to 1?

1.38 In compact-disc (CD) audio technology, the audio signal is sampled at 44.1 kHz. Each sample is represented by 16 bits. What is the speed of this system in bits per second?

Section 1.4: Amplifiers

1.39 Various amplifier and load combinations are measured as listed below using rms values. For each, find the voltage, current, and power gains (A_v, A_i, and A_p, respectively) both as ratios and in dB:

(a) $v_I = 100$ mV, $i_I = 100$ μA, $v_O = 10$ V, $R_L = 100$ Ω
(b) $v_I = 10$ μV, $i_I = 100$ nA, $v_O = 1$ V, $R_L = 10$ kΩ
(c) $v_I = 1$ V, $i_I = 1$ mA, $v_O = 5$ V, $R_L = 10$ Ω

1.40 An amplifier operating from ±3-V supplies provides a 2.2-V peak sine wave across a 100-Ω load when provided with a 0.2-V peak input from which 1.0 mA peak is drawn. The average current in each supply is measured to be 20 mA.Find the voltage gain, current gain, and power gain expressed as ratios and in decibels as well as the supply power, amplifier dissipation, and amplifier efficiency.

1.41 An amplifier using balanced power supplies is known to saturate for signals extending within 1.0 V of either supply. For linear operation, its gain is 200 V/V. What is the rms value of the largest undistorted sine-wave output available, and input needed, with ±5-V supplies? With ±10-V supplies? With ±15-V supplies?

1.42 Symmetrically saturating amplifiers, operating in the so-called clipping mode, can be used to convert sine waves to pseudo-square waves. For an amplifier with a small-signal gain of 1000 and clipping levels of ±10 V, what peak value of input sinusoid is needed to produce an output whose extremes are just at the edge of clipping? Clipped 90% of the time? Clipped 99% of the time?

Section 1.5: Circuit Models for Amplifiers

1.43 Consider the voltage-amplifier circuit model shown in Fig. 1.16(b), in which $A_{vo} = 100$ V/V under the following conditions:

(a) $R_i = 10R_s$, $R_L = 10R_o$
(b) $R_i = R_s$, $R_L = R_o$
(c) $R_i = R_s/10$, $R_L = R_o/10$

Calculate the overall voltage gain v_o/v_s in each case, expressed both directly and in decibels.

1.44 An amplifier with 40 dB of small-signal, open-circuit voltage gain, an input resistance of 1 MΩ, and an output resistance of 100 Ω, drives a load of 500 Ω. What voltage and power gains (expressed in dB) would you expect with the load connected? If the amplifier has a peak output-current limitation of 20 mA, what is the rms value of the largest sine-wave input for which an undistorted output is possible? What is the corresponding output power available?

1.45 A 10-mV signal source having an internal resistance of 100 kΩ is connected to an amplifier for which the input resistance is 10 kΩ, the open-circuit voltage gain is 1000 V/V, and the output resistance is 1 kΩ. The amplifier is connected in turn to a 100-Ω load. What overall voltage gain results as measured from the source internal voltage to the load? Where did all the gain go? What would the gain be if the source was connected directly to the load? What is the ratio of these two gains? This ratio is a useful measure of the benefit the amplifier brings.

1.46 A buffer amplifier with a gain of 1 V/V has an input resistance of 1 MΩ and an output resistance of 20 Ω. It is connected between a 1-V, 200-kΩ source and a 100-Ω

load. What load voltage results? What are the corresponding voltage, current, and power gains (in dB)?

1.47 Consider the cascade amplifier of Example 1.3. Find the overall voltage gain v_o/v_s obtained when the first and second stages are interchanged. Compare this value with the result in Example 1.3, and comment.

1.48 You are given two amplifiers, A and B, to connect in cascade between a 10-mV, 100-kΩ source and a 100-Ω load. The amplifiers have voltage gain, input resistance, and output resistance as follows: for A, 100 V/V, 100 kΩ, 10 kΩ, respectively; for B, 10 V/V, 10 kΩ, 1 kΩ, respectively. Your problem is to decide how the amplifiers should be connected. To proceed, evaluate the two possible connections between source S and load L, namely, SABL and SBAL. Find the voltage gain for each both as a ratio and in decibels. Which amplifier arrangement is best?

D *1.49 A designer has available voltage amplifiers with an input resistance of 10 kΩ, an output resistance of 1 kΩ, and an open-circuit voltage gain of 10. The signal source has a 10-kΩ resistance and provides a 5-mV rms signal, and it is required to provide a signal of at least 3 V rms to a 200-Ω load. How many amplifier stages are required? What is the output voltage actually obtained?

D *1.50 Design an amplifier that provides 0.5 W of signal power to a 100-Ω load resistance. The signal source provides a 30-mV rms signal and has a resistance of 0.5 MΩ. Three types of voltage-amplifier stages are available:

(a) A high-input-resistance type with $R_i = 1$ MΩ, $A_{vo} = 10$, and $R_o = 10$ kΩ
(b) A high-gain type with $R_i = 10$ kΩ, $A_{vo} = 100$, and $R_o = 1$ kΩ
(c) A low-output-resistance type with $R_i = 10$ kΩ, $A_{vo} = 1$, and $R_o = 20$ Ω

Design a suitable amplifier using a combination of these stages. Your design should utilize the minimum number of stages and should ensure that the signal level is not reduced below 10 mV at any point in the amplifier chain. Find the load voltage and power output realized.

D *1.51 It is required to design a voltage amplifier to be driven from a signal source having a 5-mV peak amplitude and a source resistance of 10 kΩ to supply a peak output of 2 V across a 1-kΩ load.

(a) What is the required voltage gain from the source to the load?
(b) If the peak current available from the source is 0.1 μA, what is the smallest input resistance allowed? For the design with this value of R_i, find the overall current gain and power gain.
(c) If the amplifier power supply limits the peak value of the output open-circuit voltage to 3 V, what is the largest output resistance allowed?
(d) For the design with R_i as in (b) and R_o as in (c), what is the required value of open-circuit voltage gain, i.e., $\left.\frac{v_o}{v_i}\right|_{R_L=\infty}$, of the amplifier?
(e) If, as a possible design option, you are able to increase R_i to the nearest value of the form 1×10^n Ω and to decrease R_o to the nearest value of the form 1×10^m Ω, find (i) the input resistance achievable; (ii) the output resistance achievable; and (iii) the open-circuit voltage gain now required to meet the specifications.

D 1.52 A voltage amplifier with an input resistance of 20 kΩ, an output resistance of 100 Ω, and a gain of 1000 V/V is connected between a 100-kΩ source with an open-circuit voltage of 10 mV and a 100-Ω load. For this situation:

(a) What output voltage results?
(b) What is the voltage gain from source to load?
(c) What is the voltage gain from the amplifier input to the load?
(d) If the output voltage across the load is twice that needed and there are signs of internal amplifier overload, suggest the location and value of a single resistor that would produce the desired output. Choose an arrangement that would cause minimum disruption to an operating circuit. (*Hint:* Use parallel rather than series connections.)

1.53 A voltage amplifier delivers 200 mV across a load resistance of 1 kΩ. It was found that the output voltage decreases by 5 mV when R_L is decreased to 780 Ω. What are the values of the open-circuit output voltage and the output resistance of the amplifier?

1.54 A current amplifier supplies 1 mA to a load resistance of 1 kΩ. When the load resistance is increased to 12 kΩ, the output current decreases to 0.5 mA. What are the values of the short-circuit output current and the output resistance of the amplifier?

1.55 A current amplifier for which $R_i = 100\ \Omega$, $R_o = 10\ \text{k}\Omega$, and $A_{is} = 100$ A/A is to be connected between a 100-mV source with a resistance of 10 kΩ and a load of 1 kΩ. What are the values of current gain i_o/i_i, of voltage gain v_o/v_s, and of power gain expressed directly and in decibels?

1.56 A transconductance amplifier with $R_i = 2\ \text{k}\Omega$, $G_m = 60$ mA/V, and $R_o = 20\ \text{k}\Omega$ is fed with a voltage source having a source resistance of 1 kΩ and is loaded with a 1-kΩ resistance. Find the voltage gain realized.

D **1.57 A designer is required to provide, across a 10-kΩ load, the weighted sum, $v_O = 10v_1 + 20v_2$, of input signals v_1 and v_2, each having a source resistance of 10 kΩ. She has a number of transconductance amplifiers for which the input and output resistances are both 10 kΩ and $G_m = 20$ mA/V, together with a selection of suitable resistors. Sketch an appropriate amplifier topology with additional resistors selected to provide the desired result. Your design should utilize the minimum number of amplifiers and resistors. (*Hint*: In your design, arrange to add currents.)

1.58 Figure P1.58 shows a transconductance amplifier whose output is *fed back* to its input. Find the input resistance R_{in} of the resulting one-port network. (*Hint*: Apply a test voltage v_x between the two input terminals, and find the current i_x drawn from the source. Then, $R_{in} \equiv v_x/i_x$.)

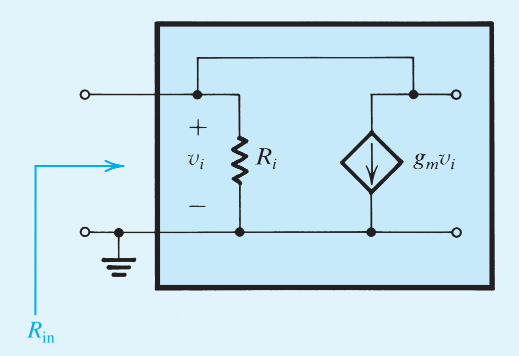


Figure P1.58

D 1.59 It is required to design an amplifier to sense the open-circuit output voltage of a transducer and to provide a proportional voltage across a load resistor. The equivalent source resistance of the transducer is specified to vary in the range of 1 kΩ to 10 kΩ. Also, the load resistance varies in the range of 1 kΩ to 10 kΩ. The change in load voltage corresponding to the specified change in R_s should be 10% at most. Similarly, the change in load voltage corresponding to the specified change in R_L should be limited to 10%. Also, corresponding to a 10-mV transducer open-circuit output voltage, the amplifier should provide a minimum of 1 V across the load. What type of amplifier is required? Sketch its circuit model, and specify the values of its parameters. Specify appropriate values for R_i and R_o of the form $1 \times 10^m\ \Omega$.

D 1.60 It is required to design an amplifier to sense the short-circuit output current of a transducer and to provide a proportional current through a load resistor. The equivalent source resistance of the transducer is specified to vary in the range of 1 kΩ to 10 kΩ. Similarly, the load resistance is known to vary over the range of 1 kΩ to 10 kΩ. The change in load current corresponding to the specified change in R_s is required to be limited to 10%. Similarly, the change in load current corresponding to the specified change in R_L

should be 10% at most. Also, for a nominal short-circuit output current of the transducer of 10 μA, the amplifier is required to provide a minimum of 1 mA through the load. What type of amplifier is required? Sketch the circuit model of the amplifier, and specify values for its parameters. Select appropriate values for R_i and R_o in the form $1 \times 10^m\ \Omega$.

D 1.61 It is required to design an amplifier to sense the open-circuit output voltage of a transducer and to provide a proportional current through a load resistor. The equivalent source resistance of the transducer is specified to vary in the range of 1 kΩ to 10 kΩ. Also, the load resistance is known to vary in the range of 1 kΩ to 10 kΩ. The change in the current supplied to the load corresponding to the specified change in R_s is to be 10% at most. Similarly, the change in load current corresponding to the specified change in R_L is to be 10% at most. Also, for a nominal transducer open-circuit output voltage of 10 mV, the amplifier is required to provide a minimum of 1 mA current through the load. What type of amplifier is required? Sketch the amplifier circuit model, and specify values for its parameters. For R_i and R_o, specify values in the form $1 \times 10^m\ \Omega$.

D 1.62 It is required to design an amplifier to sense the short-circuit output current of a transducer and to provide a proportional voltage across a load resistor. The equivalent source resistance of the transducer is specified to vary in the range of 1 kΩ to 10 kΩ. Similarly, the load resistance is known to vary in the range of 1 kΩ to 10 kΩ. The change in load voltage corresponding to the specified change in R_s should be 10% at most. Similarly, the change in load voltage corresponding to the specified change in R_L is to be limited to 10%. Also, for a nominal transducer short-circuit output current of 10 μA, the amplifier is required to provide a minimum voltage across the load of 1 V. What type of amplifier is required? Sketch its circuit model, and specify the values of the model parameters. For R_i and R_o, specify appropriate values in the form $1 \times 10^m\ \Omega$.

1.63 For the circuit in Fig. P1.63, show that

$$\frac{v_c}{v_b} = \frac{-\beta R_L}{r_\pi + (\beta + 1)R_E}$$

and

$$\frac{v_e}{v_b} = \frac{R_E}{R_E + [r_\pi/(\beta + 1)]}$$

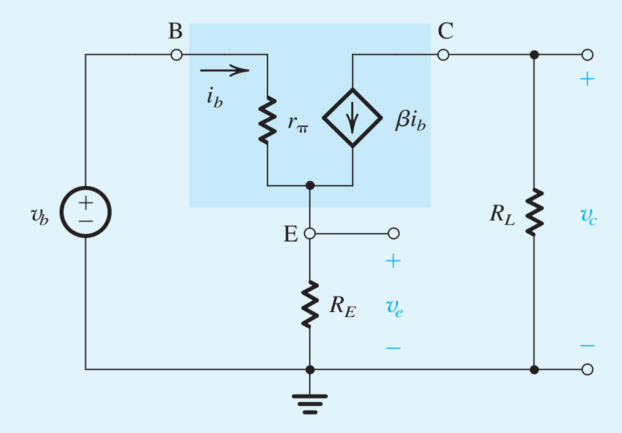


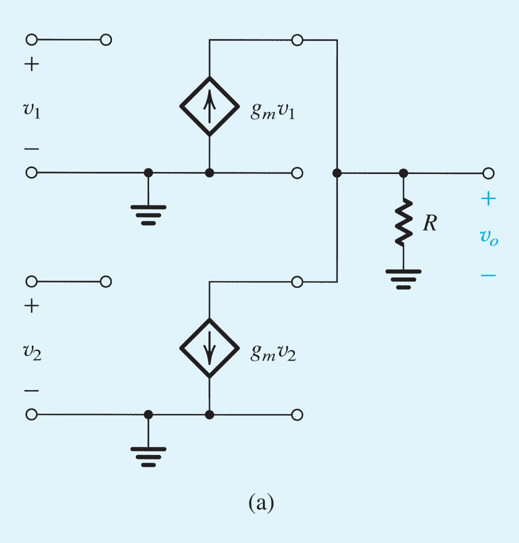
Figure P1.63

1.64 An amplifier with an input resistance of 5 kΩ, when driven by a current source of 1 μA and a source resistance of 200 kΩ, has a short-circuit output current of 5 mA and an open-circuit output voltage of 10 V. If the amplifier is used to drive a 2-kΩ load, give the values of the voltage gain, current gain, and power gain expressed as ratios and in decibels.

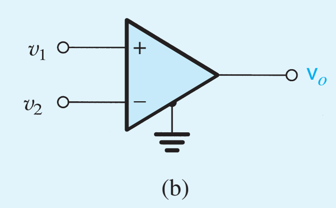
1.65 Figure P1.65(a) shows two transconductance amplifiers connected in a special configuration. Find v_o in terms of v_1 and v_2. Let $g_m = 100$ mA/V and $R = 5$ kΩ. If $v_1 = v_2 = 1$ V, find the value of v_o. Also, find v_o for the case $v_1 = 1.01$ V and $v_2 = 0.99$ V. (*Note*: This circuit is called a **differential amplifier** and is given the symbol shown

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

in Fig. P1.65(b). A particular type of differential amplifier known as an **operational amplifier** will be studied in Chapter 2.)



(a)



(b)

Figure P1.65

1.66 Any linear two-port network including linear amplifiers can be represented by one of four possible parameter sets, given in Appendix C. For the voltage amplifier, the most convenient representation is in terms of the g parameters. If the amplifier input port is labeled as port 1 and the output port as port 2, its g-parameter representation is described by the two equations:

$$I_1 = g_{11}V_1 + g_{12}I_2$$

$$V_2 = g_{21}V_1 + g_{22}I_2$$

Figure P1.66 shows an equivalent circuit representation of these two equations. By comparing this equivalent circuit to that of the voltage amplifier in Fig. 1.16(a), identify corresponding currents and voltages as well as the correspondence between the parameters of the amplifier equivalent circuit and the g parameters. Hence give the g parameter that corresponds to each of R_i, A_{vo}, and R_o. Notice that there is an additional g parameter with no correspondence in the amplifier equivalent circuit. Which one? What does it signify? What assumption did we make about the amplifier that resulted in the absence of this particular g parameter from the equivalent circuit in Fig. 1.16(a)?

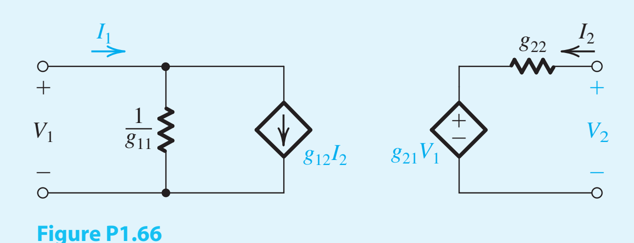


Figure P1.66

Section 1.6: Frequency Response of Amplifiers

1.67 Use the voltage-divider rule to derive the transfer functions $T(s) \equiv V_o(s)/V_i(s)$ of the circuits shown in Fig. 1.22, and show that the transfer functions are of the form given at the top of Table 1.2.

1.68 Figure P1.68 shows a signal source connected to the input of an amplifier. Here R_s is the source resistance, and R_i and C_i are the input resistance and input capacitance, respectively, of the amplifier. Derive an expression for $V_i(s)/V_s(s)$, and show that it is of the low-pass STC type. Find the 3-dB frequency for the case $R_s = 10\text{ k}\Omega$, $R_i = 40\text{ k}\Omega$, and $C_i = 5\text{ pF}$.

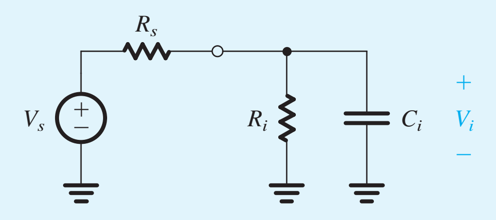


Figure P1.68

1.69 For the circuit shown in Fig. P1.69, find the transfer function $T(s) = V_o(s)/V_i(s)$, and arrange it in the appropriate standard form from Table 1.2. Is this a high-pass or a low-pass network? What is its transmission at very high frequencies? [Estimate this directly, as well as by letting $s \to \infty$ in your expression for $T(s)$.] What is the corner frequency ω_0? For $R_1 = 10\ \text{k}\Omega$, $R_2 = 40\ \text{k}\Omega$, and $C = 1\ \mu\text{F}$, find f_0. What is the value of $|T(j\omega_o)|$?

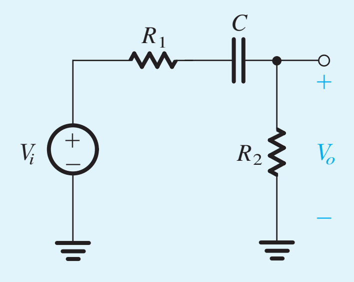


Figure P1.69

D 1.70 It is required to couple a voltage source V_s with a resistance R_s to a load R_L via a capacitor C. Derive an expression for the transfer function from source to load (i.e., V_l/V_s), and show that it is of the high-pass STC type. For $R_s = 5\ \text{k}\Omega$ and $R_L = 20\ \text{k}\Omega$, find the smallest coupling capacitor that will result in a 3-dB frequency no greater than 100 Hz.

1.71 Measurement of the frequency response of an amplifier yields the data in the following table:

f **(Hz)**	**|*T*| (dB)**	**∠*T*** **(°)**
0	40	0
100	40	0
1000		
10^4	37	−45
10^5	20	
	0	

Provide plausible approximate values for the missing entries. Also, sketch and clearly label the magnitude frequency response (i.e., provide a Bode plot) for this amplifier.

1.72 Measurement of the frequency response of an amplifier yields the data in the following table:

f **(Hz)**		10	10^2	10^3	10^4	10^5	10^6	10^7	
|*T*| (dB)	0	20	37	40			37	20	0

Provide approximate plausible values for the missing table entries. Also, sketch and clearly label the magnitude frequency response (Bode plot) of this amplifier.

1.73 The unity-gain voltage amplifiers in the circuit of Fig. P1.73 have infinite input resistances and zero output

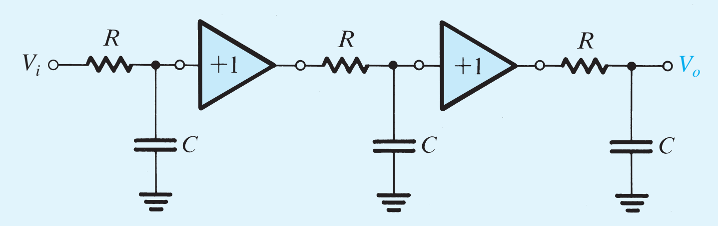


Figure P1.73

resistances and thus function as perfect buffers. Furthermore, assume that their gain is frequency independent. Convince yourself that the overall gain V_o/V_i will drop by 3 dB below the value at dc at the frequency for which the gain of each *RC* circuit is 1.0 dB down. What is that frequency in terms of *CR*?

1.74 A manufacturing error causes an internal node of a high-frequency amplifier whose Thévenin-equivalent node resistance is 100 kΩ to be accidentally shunted to ground by a capacitor (i.e., the node is connected to ground through a capacitor). If the measured 3-dB bandwidth of the amplifier is reduced from the expected 5 MHz to 100 kHz, estimate the value of the shunting capacitor. If the original cutoff frequency can be attributed to a small parasitic capacitor at the same internal node (i.e., between the node and ground), what would you estimate it to be?

D *1.75 A designer wishing to lower the overall upper 3-dB frequency of a three-stage amplifier to 10 kHz considers shunting one of two nodes: Node A, between the output of the first stage and the input of the second stage, and Node B, between the output of the second stage and the input of the third stage, to ground with a small capacitor. While measuring the overall frequency response of the amplifier, she connects a capacitor of 1 nF, first to node A and then to node B, lowering the 3-dB frequency from 3 MHz to 200 kHz and 20 kHz, respectively. If she knows that each amplifier stage has an input resistance of 100 kΩ, what output resistance must the driving stage have at node A? At node B? What capacitor value should she connect to which node to solve her design problem most economically?

D 1.76 An amplifier with an input resistance of 100 kΩ and an output resistance of 1 kΩ is to be capacitor-coupled to a 10-kΩ source and a 1-kΩ load. Available capacitors have values only of the form 1×10^{-n} F. What are the values of the smallest capacitors needed to ensure that the corner frequency associated with each is less than 100 Hz? What actual corner frequencies result? For the situation in which the basic amplifier has an open-circuit voltage gain (A_{vo}) of 100 V/V, find an expression for $T(s) = V_o(s)/V_s(s)$.

***1.77** A voltage amplifier has the transfer function

$$A_v = \frac{1000}{\left(1 + j\dfrac{f}{10^5}\right)\left(1 + \dfrac{10^2}{jf}\right)}$$

Using the Bode plots for low-pass and high-pass STC networks (Figs. 1.23 and 1.24), sketch a Bode plot for $|A_v|$. Give approximate values for the gain magnitude at $f = 10$ Hz, 10^2 Hz, 10^3 Hz, 10^4 Hz, 10^5 Hz, 10^6 Hz, 10^7 Hz, and 10^8 Hz. Find the bandwidth of the amplifier (defined as the frequency range over which the gain remains within 3 dB of the maximum value).

***1.78** For the circuit shown in Fig. P1.78, first evaluate $T_i(s) = V_i(s)/V_s(s)$ and the corresponding cutoff (corner) frequency. Second, evaluate $T_o(s) = V_o(s)/V_i(s)$ and the corresponding cutoff frequency. Put each of the transfer functions in the standard form (see Table 1.2), and combine them to form the overall transfer function, $T(s) = T_i(s) \times T_o(s)$. Provide a Bode magnitude plot for $|T(j\omega)|$. What is the bandwidth between 3-dB cutoff points?

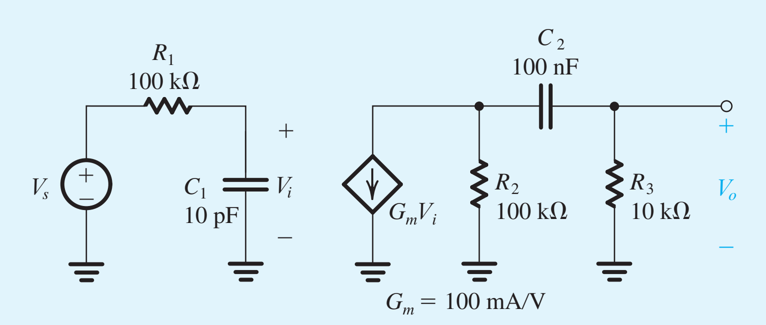


Figure P1.78

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

D **1.79 A transconductance amplifier having the equivalent circuit shown in Table 1.1 is fed with a voltage source V_s having a source resistance R_s, and its output is connected to a load consisting of a resistance R_L in parallel with a capacitance C_L. For given values of R_s, R_L, and C_L, it is required to specify the values of the amplifier parameters R_i, G_m, and R_o to meet the following design constraints:

(a) At most, $x\%$ of the input signal is lost in coupling the signal source to the amplifier (i.e., $V_i \geq [1 - (x/100)]V_s$).

(b) The 3-dB frequency of the amplifier is equal to or greater than a specified value $f_{3\text{ dB}}$.

(c) The dc gain V_o/V_s is equal to or greater than a specified value A_0.

Show that these constraints can be met by selecting

$$R_i \geq \left(\frac{100}{x} - 1\right)R_s$$

$$R_o \leq \frac{1}{2\pi f_{3\text{dB}} C_L - (1/R_L)}$$

$$G_m \geq \frac{A_0/[1-(x/100)]}{(R_L \parallel R_o)}$$

Find R_i, R_o, and G_m for $R_s = 10\text{ k}\Omega$, $x = 10\%$, $A_0 = 100$ V/V, $R_L = 10\text{ k}\Omega$, $C_L = 20$ pF, and $f_{3\text{ dB}} = 2$ MHz.

***1.80** Use the voltage-divider rule to find the transfer function $V_o(s)/V_i(s)$ of the circuit in Fig. P1.80. Show that the transfer function can be made independent of frequency if the condition $C_1R_1 = C_2R_2$ applies. Under this condition the circuit is called a **compensated attenuator** and is frequently employed in the design of oscilloscope probes. Find the transmission of the compensated attenuator in terms of R_1 and R_2.

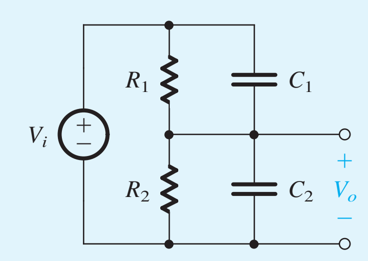


Figure P1.80

***1.81** An amplifier with a frequency response of the type shown in Fig. 1.21 is specified to have a phase shift of magnitude no greater than 5.7° over the amplifier bandwidth, which extends from 100 Hz to 1 kHz. It has been found that the gain falloff at the low-frequency end is determined by the response of a high-pass STC circuit and that at the high-frequency end it is determined by a low-pass STC circuit. What do you expect the corner frequencies of these two circuits to be? What is the drop in gain in decibels (relative to the maximum gain) at the two frequencies that define the amplifier bandwidth? What are the frequencies at which the drop in gain is 3 dB? (*Hint*: Refer to Figs. 1.23 and 1.24.)

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

CHAPTER 2

Operational Amplifiers

IN THIS CHAPTER YOU WILL LEARN

1. The terminal characteristics of the ideal op amp.
2. How to analyze circuits containing op amps, resistors, and capacitors.
3. How to use op amps to design amplifiers having precise characteristics.
4. How to design more sophisticated op-amp circuits, including summing amplifiers, instrumentation amplifiers, integrators, and differentiators.
5. Important nonideal characteristics of op amps and how these limit the performance of basic op-amp circuits.

Introduction

Having learned basic amplifier concepts and terminology, we are now ready to undertake the study of a circuit building block of universal importance: the operational amplifier (op amp). Op amps have been in use for a long time, their initial applications being primarily in the areas of analog computation and sophisticated instrumentation. Early op amps were constructed from discrete components (vacuum tubes and then transistors, and resistors), and their cost was prohibitively high (tens of dollars). In the mid-1960s the first integrated-circuit (IC) op amp was produced. This unit (the μA 709) was made up of a relatively large number of transistors and resistors all on the same silicon chip. Although its characteristics were poor (by today's standards) and its price was still quite high, its appearance signaled a new era in electronic circuit design. Electronics engineers started using op amps in large quantities, which caused their price to drop dramatically. They also demanded better-quality op amps. Semiconductor manufacturers responded quickly, and within the span of a few years, high-quality op amps became available at extremely low prices (tens of cents) from a large number of suppliers.

One of the reasons for the popularity of the op amp is its versatility. As we will shortly see, one can do almost anything with op amps! Equally important is the fact that the IC op amp has characteristics that closely approach the assumed ideal. This implies that it is quite easy to design circuits using the IC op amp. Also, op-amp circuits work at performance levels that are quite close to those predicted theoretically. It is for this reason that we are studying op amps at this early stage. It is expected that by the end of this chapter the reader should be able to successfully design nontrivial circuits using op amps.

As already implied, an IC op amp is made up of a large number (about 20) of transistors together with resistors, and (usually) one capacitor connected in a rather complex circuit. Since

we have not yet studied transistor circuits, the circuit inside the op amp will not be discussed in this chapter. Rather, we will treat the op amp as a circuit building block and study its terminal characteristics and its applications. This approach is quite satisfactory in many op-amp applications. Nevertheless, for the more difficult and demanding applications it is quite useful to know what is inside the op-amp package. This topic will be studied in Chapter 13. More advanced applications of op amps will appear in later chapters.

2.1 The Ideal Op Amp

2.1.1 The Op-Amp Terminals

From a signal point of view the op amp has three terminals: two input terminals and one output terminal. Figure 2.1 shows the symbol we shall use to represent the op amp. Terminals 1 and 2 are input terminals, and terminal 3 is the output terminal. As explained in Section 1.4, amplifiers require dc power to operate. Most IC op amps require two dc power supplies, as shown in Fig. 2.2. Two terminals, 4 and 5, are brought out of the op-amp package and connected to a positive voltage V_{CC} and a negative voltage $-V_{EE}$, respectively. In Fig. 2.2(b) we explicitly show the two dc power supplies as batteries with a common ground. It is interesting to note that the reference grounding point in op-amp circuits is just the common terminal of the two power supplies; that is, no terminal of the op-amp package is physically connected to ground. In what follows we will not, for simplicity, explicitly show the op-amp power supplies.

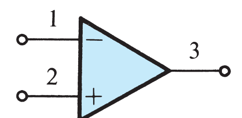


Figure 2.1 Circuit symbol for the op amp.

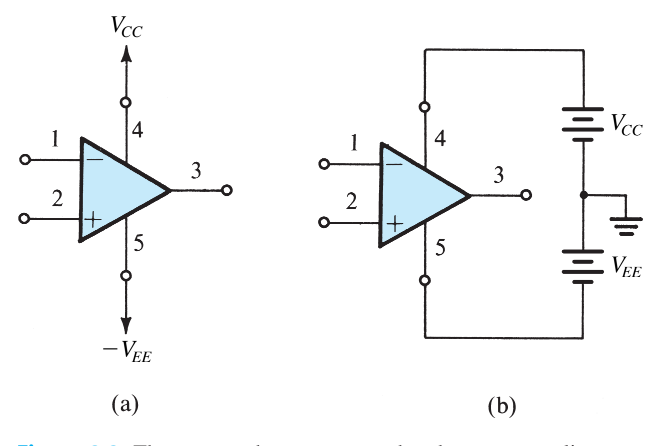


Figure 2.2 The op amp shown connected to dc power supplies.

In addition to the three signal terminals and the two power-supply terminals, an op amp may have other terminals for specific purposes. These other terminals can include terminals for frequency compensation and terminals for offset nulling; both functions will be explained in later sections.

EXERCISE

2.1 What is the minimum number of terminals required by a single op amp? What is the minimum number of terminals required on an integrated-circuit package containing four op amps (called a quad op amp)?
Ans. 5; 14

2.1.2 Function and Characteristics of the Ideal Op Amp

We now consider the circuit function of the op amp. The op amp is designed to sense the difference between the voltage signals applied at its two input terminals (i.e., the quantity $v_2 - v_1$), multiply this by a number A, and cause the resulting voltage $A(v_2 - v_1)$ to appear at output terminal 3. Thus $v_3 = A(v_2 - v_1)$. Here it should be emphasized that when we talk about the voltage at a terminal we mean the voltage between that terminal and ground; thus v_1 means the voltage applied between terminal 1 and ground.

The ideal op amp is not supposed to draw any input current; that is, the signal current into terminal 1 and the signal current into terminal 2 are both zero. In other words, *the input impedance of an ideal op amp is supposed to be infinite.*

How about the output terminal 3? This terminal is supposed to act as the output terminal of an ideal voltage source. That is, the voltage between terminal 3 and ground will always be equal to $A(v_2 - v_1)$, independent of the current that may be drawn from terminal 3 into a load impedance. In other words, *the output impedance of an ideal op amp is supposed to be zero.*

Putting together all of the above, we arrive at the equivalent circuit model shown in Fig. 2.3. Note that the output is in phase with (has the same sign as) v_2 and is out of phase with (has the opposite sign of) v_1. For this reason, input terminal 1 is called the **inverting input terminal** and is distinguished by a "−" sign, while input terminal 2 is called the **noninverting input terminal** and is distinguished by a "+" sign.

As can be seen from the above description, the op amp responds only to the *difference* signal $v_2 - v_1$ and hence ignores any signal *common* to both inputs. That is, if $v_1 = v_2 = 1$ V, then the output will (ideally) be zero. We call this property **common-mode rejection**, and we conclude that an ideal op amp has zero common-mode gain or, equivalently, infinite common-mode rejection. We will have more to say about this point later. For the time being note that the op amp is a **differential-input, single-ended-output** amplifier, with the latter term referring to the fact that the output appears between terminal 3 and ground.[1]

[1]Some op amps are designed to have differential outputs. This topic will not be discussed in this book. Rather, we confine ourselves here to single-ended-output op amps, which constitute the vast majority of commercially available op amps.

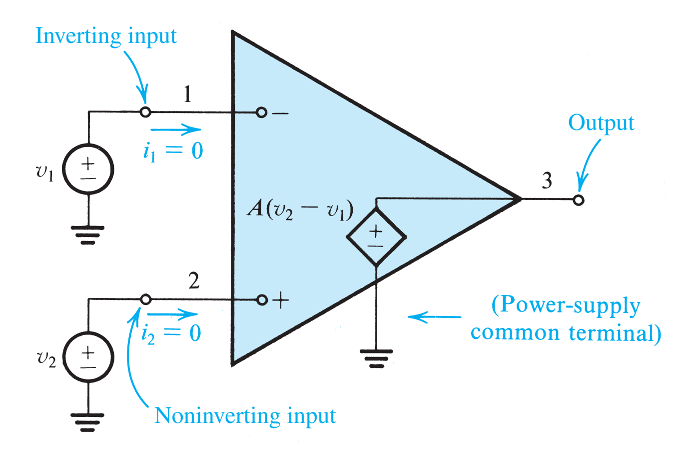


Figure 2.3 Equivalent circuit of the ideal op amp.

Furthermore, gain A is called the **differential gain**, for obvious reasons. Perhaps not so obvious is another name that we will attach to A: the **open-loop gain**. The reason for this name will become obvious later on when we "close the loop" around the op amp and define another gain, the closed-loop gain.

An important characteristic of op amps is that they are **direct-coupled** or **dc amplifiers**, where dc stands for direct-coupled (it could equally well stand for direct current, since a direct-coupled amplifier is one that amplifies signals whose frequency is as low as zero). The fact that op amps are direct-coupled devices will allow us to use them in many important applications. Unfortunately, though, the direct-coupling property can cause some serious practical problems, as will be discussed in a later section.

How about bandwidth? The ideal op amp has a gain A that remains constant down to zero frequency and up to infinite frequency. That is, ideal op amps will amplify signals of any frequency with equal gain, and are thus said to have *infinite bandwidth.*

We have discussed all of the properties of the ideal op amp except for one, which in fact is the most important. This has to do with the value of A. *The ideal op amp should have a gain A whose value is very large and ideally infinite*. One may justifiably ask: If the gain A is infinite, how are we going to use the op amp? The answer is very simple: In almost all applications the op amp will *not* be used alone in a so-called open-loop configuration. Rather, we will use other components to apply feedback to close the loop around the op amp, as will be illustrated in detail in Section 2.2.

For future reference, Table 2.1 lists the characteristics of the ideal op amp.

Table 2.1 Characteristics of the Ideal Op Amp

1. Infinite input impedance
2. Zero output impedance
3. Zero common-mode gain or, equivalently, infinite common-mode rejection
4. Infinite open-loop gain A
5. Infinite bandwidth

2.1.3 Differential and Common-Mode Signals

The differential input signal v_{Id} is simply the difference between the two input signals v_1 and v_2; that is,

$$v_{Id} = v_2 - v_1 \qquad (2.1)$$

The common-mode input signal v_{Icm} is the average of the two input signals v_1 and v_2; namely,

$$v_{Icm} = \tfrac{1}{2}(v_1 + v_2) \qquad (2.2)$$

Equations (2.1) and (2.2) can be used to express the input signals v_1 and v_2 in terms of their differential and common-mode components as follows:

$$v_1 = v_{Icm} - v_{Id}/2 \qquad (2.3)$$

and

$$v_2 = v_{Icm} + v_{Id}/2 \qquad (2.4)$$

These equations can in turn lead to the pictorial representation in Fig. 2.4.

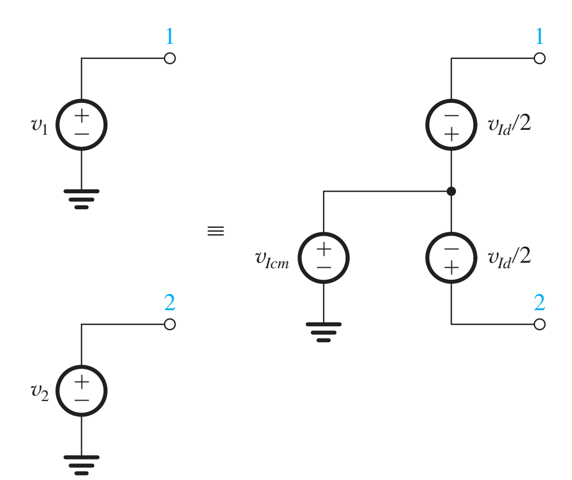


Figure 2.4 Representation of the signal sources v_1 and v_2 in terms of their differential and common-mode components.

EXERCISES

2.2 Consider an op amp that is ideal except that its open-loop gain $A = 10^3$. The op amp is used in a feedback circuit, and the voltages appearing at two of its three signal terminals are measured. In each of the following cases, use the measured values to find the expected value of the voltage at the third terminal. Also give the differential and common-mode input signals in each case. (a) $v_2 = 0$ V and $v_3 = 2$ V; (b) $v_2 = +5$ V and $v_3 = -10$ V; (c) $v_1 = 1.002$ V and $v_2 = 0.998$ V; (d) $v_1 = -3.6$ V and $v_3 = -3.6$ V.
Ans. (a) $v_1 = -0.002$ V, $v_{Id} = 2$ mV, $v_{Icm} = -1$ mV; (b) $v_1 = +5.01$ V, $v_{Id} = -10$ mV, $v_{Icm} = 5.005 \simeq 5$ V; (c) $v_3 = -4$ V, $v_{Id} = -4$ mV, $v_{Icm} = 1$ V; (d) $v_2 = -3.6036$ V, $v_{Id} = -3.6$ mV, $v_{Icm} \simeq -3.6$ V

2.3 The internal circuit of a particular op amp can be modeled by the circuit shown in Fig. E2.3. Express v_3 as a function of v_1 and v_2. For the case $G_m = 10$ mA/V, $R = 10$ kΩ, and $\mu = 100$, find the value of the open-loop gain A.

Ans. $v_3 = \mu G_m R(v_2 - v_1)$; $A = 10{,}000$ V/V or 80 dB

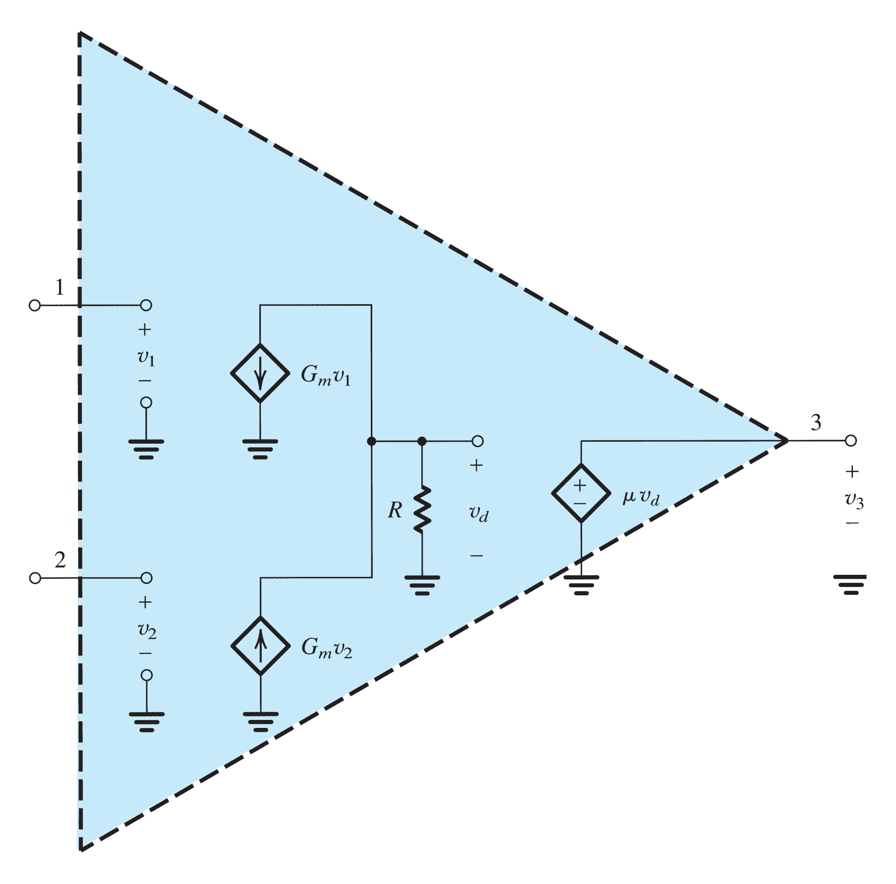


Figure E2.3

2.2 The Inverting Configuration

As mentioned above, op amps are not used alone; rather, the op amp is connected to passive components in a feedback circuit. There are two such basic circuit configurations employing an op amp and two resistors: the inverting configuration, which is studied in this section, and the noninverting configuration, which we shall study in the next section.

Figure 2.5 shows the inverting configuration. It consists of one op amp and two resistors R_1 and R_2. Resistor R_2 is connected from the output terminal of the op amp, terminal 3, *back* to the *inverting* or *negative* input terminal, terminal 1. We speak of R_2 as applying **negative feedback**; if R_2 were connected between terminals 3 and 2 we would have called this **positive feedback**. Note also that R_2 *closes the loop* around the op amp. In addition to adding R_2, we have grounded terminal 2 and connected a resistor R_1 between terminal 1 and an input signal source

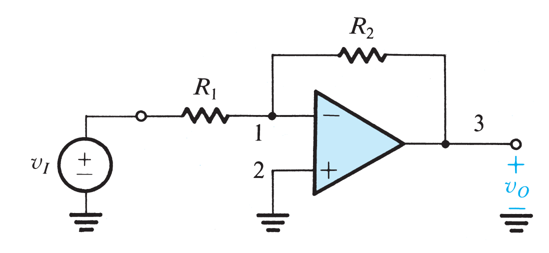


Figure 2.5 The inverting closed-loop configuration.

with a voltage v_I. The output of the overall circuit is taken at terminal 3 (i.e., between terminal 3 and ground). Terminal 3 is, of course, a convenient point from which to take the output, since the impedance level there is ideally zero. Thus the voltage v_O will not depend on the value of the current that might be supplied to a load impedance connected between terminal 3 and ground.

2.2.1 The Closed-Loop Gain

We now wish to analyze the circuit in Fig. 2.5 to determine the **closed-loop gain** G, defined as

$$G \equiv \frac{v_O}{v_I}$$

We will do so assuming the op amp to be ideal. Figure 2.6(a) shows the equivalent circuit, and the analysis proceeds as follows: The gain A is very large (ideally infinite). If we assume that the circuit is "working" and producing a finite output voltage at terminal 3, then the voltage between the op-amp input terminals should be negligibly small and ideally zero. Specifically, if we call the output voltage v_O, then, by definition,

$$v_2 - v_1 = \frac{v_O}{A} = 0$$

It follows that the voltage at the inverting input terminal (v_1) is given by $v_1 = v_2$. That is, because the gain A approaches infinity, the voltage v_1 approaches and ideally equals v_2. We speak of this as the two input terminals "tracking each other in potential." We also speak of a "virtual short circuit" that exists between the two input terminals. Here the word *virtual* should be emphasized, and one should *not* make the mistake of physically shorting terminals 1 and 2 together while analyzing a circuit. A **virtual short circuit** means that whatever voltage is at 2 will automatically appear at 1 because of the infinite gain A. But terminal 2 happens to be connected to ground; thus $v_2 = 0$ and $v_1 = 0$. We speak of terminal 1 as being a **virtual ground**—that is, having zero voltage but not physically connected to ground.

Now that we have determined v_1 we are in a position to apply Ohm's law and find the current i_1 through R_1 (see Fig. 2.6) as follows:

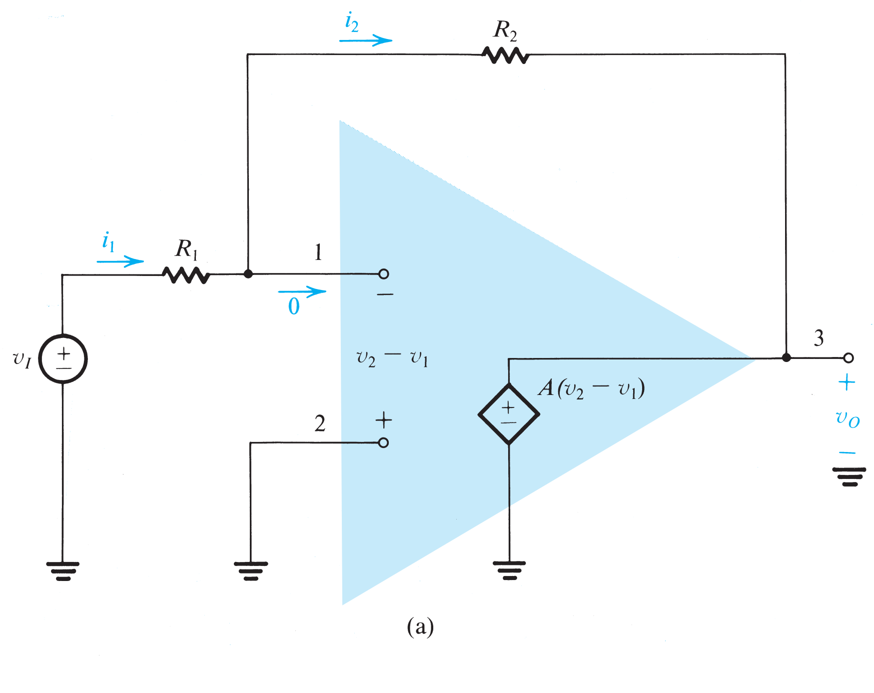
$$i_1 = \frac{v_I - v_1}{R_1} = \frac{v_I - 0}{R_1} = \frac{v_I}{R_1}$$

Where will this current go? It cannot go into the op amp, since the ideal op amp has an infinite input impedance and hence draws zero current. It follows that i_1 will have to flow through R_2 to the low-impedance terminal 3. We can then apply Ohm's law to R_2 and determine v_O; that is,

$$\begin{aligned} v_O &= v_1 - i_1 R_2 \\ &= 0 - \frac{v_I}{R_1} R_2 \end{aligned}$$

Thus,

$$\frac{v_O}{v_I} = -\frac{R_2}{R_1}$$



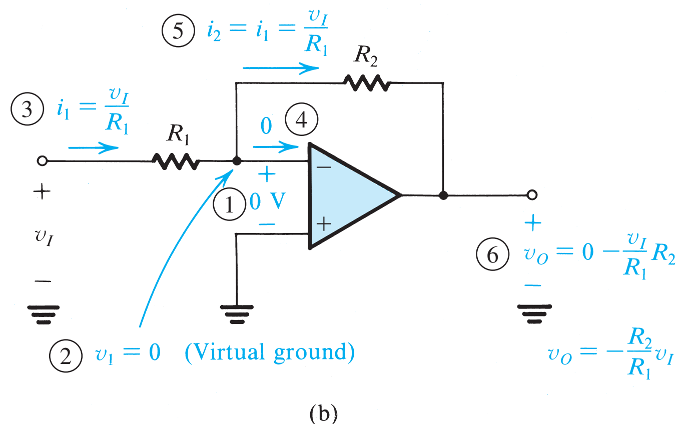


Figure 2.6 Analysis of the inverting configuration. The circled numbers indicate the order of the analysis steps.

which is the required closed-loop gain. Figure 2.6(b) illustrates these steps and indicates by the circled numbers the order in which the analysis is performed.

We thus see that the closed-loop gain is simply the ratio of the two resistances R_2 and R_1. The minus sign means that the closed-loop amplifier provides signal inversion. Thus if $R_2/R_1 = 10$ and we apply at the input (v_I) a sine-wave signal of 1 V peak-to-peak, then the output v_O will be a sine wave of 10 V peak-to-peak and phase-shifted 180° with respect to the input sine wave. Because of the minus sign associated with the closed-loop gain, this configuration is called the **inverting configuration**.

The fact that the closed-loop gain depends entirely on external passive components (resistors R_1 and R_2) is very significant. It means that we can make the closed-loop gain as accurate as we want by selecting passive components of appropriate accuracy. It also means that the closed-loop gain is (ideally) independent of the op-amp gain. This is a dramatic illustration of negative feedback: We started out with an amplifier having very large gain A, and through applying negative feedback we have obtained a closed-loop gain R_2/R_1 that is much smaller than A but is stable and predictable. That is, we are trading gain for accuracy.

2.2.2 Effect of Finite Open-Loop Gain

The points just made are more clearly illustrated by deriving an expression for the closed-loop gain under the assumption that the op-amp open-loop gain A is finite. Figure 2.7 shows the analysis. If we denote the output voltage v_O, then the voltage between the two input terminals of the op amp will be v_O/A. Since the positive input terminal is grounded, the voltage at the negative input terminal must be $-v_O/A$. The current i_1 through R_1 can now be found from

$$i_1 = \frac{v_I - (-v_O/A)}{R_1} = \frac{v_I + v_O/A}{R_1}$$

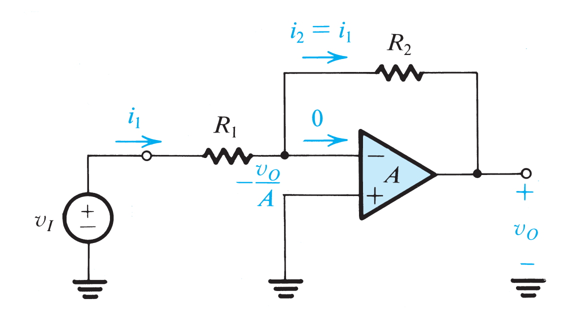


Figure 2.7 Analysis of the inverting configuration taking into account the finite open-loop gain of the op amp.

The infinite input impedance of the op amp forces the current i_1 to flow entirely through R_2. The output voltage v_O can thus be determined from

$$v_O = -\frac{v_O}{A} - i_1 R_2$$
$$= -\frac{v_O}{A} - \left(\frac{v_I + v_O/A}{R_1}\right)R_2$$

Collecting terms, the closed-loop gain G is found as

$$G \equiv \frac{v_O}{v_I} = \frac{-R_2/R_1}{1 + (1 + R_2/R_1)/A} \tag{2.5}$$

We note that as A approaches ∞, G approaches the ideal value of $-R_2/R_1$. Also, from Fig. 2.7 we see that as A approaches ∞, the voltage at the inverting input terminal approaches zero. This is the virtual-ground assumption we used in our earlier analysis when the op amp was

assumed to be ideal. Finally, note that Eq. (2.5) in fact indicates that to minimize the dependence of the closed-loop gain G on the value of the open-loop gain A, we should make

$$1 + \frac{R_2}{R_1} \ll A$$

Example 2.1

Consider the inverting configuration with $R_1 = 1\ \text{k}\Omega$ and $R_2 = 100\ \text{k}\Omega$, that is, having an ideal closed-loop gain of -100.

(a) Find the closed-loop gain for the cases $A = 10^3, 10^4$, and 10^5. In each case determine the percentage error in the magnitude of G relative to the ideal value of R_2/R_1 (obtained with $A = \infty$). Also determine the voltage v_1 that appears at the inverting input terminal when $v_I = 0.1$ V.

(b) If the open-loop gain A changes from 100,000 to 50,000 (i.e., drops by 50%), what is the corresponding percentage change in the magnitude of the closed-loop gain G?

Solution

(a) Substituting the given values in Eq. (2.5), we obtain the values given in the following table, where the percentage error ϵ is defined as

$$\epsilon \equiv \frac{|G| - (R_2/R_1)}{(R_2/R_1)} \times 100$$

The values of v_1 are obtained from $v_1 = -v_O/A = Gv_I/A$ with $v_I = -0.1$ V.

A	$\lvert G\rvert$	ϵ	v_1
10^3	90.83	−9.17%	−9.08 mV
10^4	99.00	−1.00%	−0.99 mV
10^5	99.90	−0.10%	−0.10 mV

(b) Using Eq. (2.5), we find that for $A = 50{,}000$, $|G| = 99.80$. Thus a -50% change in the open-loop gain results in a change in $|G|$ from 99.90 to 99.80, which is only -0.1%!

2.2.3 Input and Output Resistances

Assuming an ideal op amp with infinite open-loop gain, the input resistance of the closed-loop inverting amplifier of Fig. 2.5 is simply equal to R_1. This can be seen from Fig. 2.6(b), where



$$R_i \equiv \frac{v_I}{i_1} = \frac{v_I}{v_I/R_1} = R_1$$

Now recall that in Section 1.5 we learned that the amplifier input resistance forms a voltage divider with the resistance of the source that feeds the amplifier. Thus, to avoid the loss of signal strength, voltage amplifiers are required to have high input resistance. In the case of the inverting op-amp configuration we are studying, to make R_i high we should select a high value for R_1. However, if the required gain R_2/R_1 is also high, then R_2 could become impractically large

(e.g., greater than a few megohms). We may conclude that the inverting configuration suffers from a low input resistance. A solution to this problem is discussed in Example 2.2 below.

Since the output of the inverting configuration is taken at the terminals of the ideal voltage source $A(v_2 - v_1)$ (see Fig. 2.6a), it follows that the output resistance of the closed-loop amplifier is zero.

Example 2.2

Assuming the op amp to be ideal, derive an expression for the closed-loop gain v_O/v_I of the circuit shown in Fig. 2.8. Use this circuit to design an inverting amplifier with a gain of 100 and an input resistance of 1 MΩ. Assume that for practical reasons it is required not to use resistors greater than 1 MΩ. Compare your design with that based on the inverting configuration of Fig. 2.5.

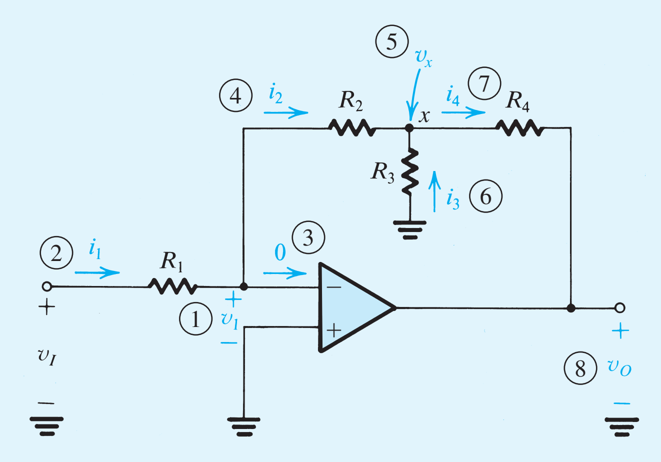


Figure 2.8 Circuit for Example 2.2. The circled numbers indicate the sequence of the steps in the analysis.

Solution

The analysis begins at the inverting input terminal of the op amp, where the voltage is

$$v_1 = \frac{-v_O}{A} = \frac{-v_O}{\infty} = 0$$

Here we have assumed that the circuit is "working" and producing a finite output voltage v_O. Knowing v_1, we can determine the current i_1 as follows:

$$i_1 = \frac{v_I - v_1}{R_1} = \frac{v_I - 0}{R_1} = \frac{v_I}{R_1}$$

Since zero current flows into the inverting input terminal, all of i_1 will flow through R_2, and thus

$$i_2 = i_1 = \frac{v_I}{R_1}$$

Now we can determine the voltage at node x:

$$v_x = v_1 - i_2R_2 = 0 - \frac{v_I}{R_1}R_2 = -\frac{R_2}{R_1}v_I$$

Example 2.2 *continued*

This in turn enables us to find the current i_3:

$$i_3 = \frac{0 - v_x}{R_3} = \frac{R_2}{R_1 R_3} v_I$$

Next, a node equation at x yields i_4:

$$i_4 = i_2 + i_3 = \frac{v_I}{R_1} + \frac{R_2}{R_1 R_3} v_I$$

Finally, we can determine v_O from

$$v_O = v_x - i_4 R_4$$

$$= -\frac{R_2}{R_1} v_I - \left(\frac{v_I}{R_1} + \frac{R_2}{R_1 R_3} v_I \right) R_4$$

Thus the voltage gain is given by

$$\frac{v_O}{v_I} = -\left[\frac{R_2}{R_1} + \frac{R_4}{R_1}\left(1 + \frac{R_2}{R_3}\right)\right]$$

which can be written in the form

$$\frac{v_O}{v_I} = -\frac{R_2}{R_1}\left(1 + \frac{R_4}{R_2} + \frac{R_4}{R_3}\right)$$

Now, since an input resistance of 1 MΩ is required, we select $R_1 = 1$ MΩ. Then, with the limitation of using resistors no greater than 1 MΩ, the maximum value possible for the first factor in the gain expression is 1 and is obtained by selecting $R_2 = 1$ MΩ. To obtain a gain of -100, R_3 and R_4 must be selected so that the second factor in the gain expression is 100. If we select the maximum allowed (in this example) value of 1 MΩ for R_4, then the required value of R_3 can be calculated to be 10.2 kΩ. Thus this circuit utilizes three 1-MΩ resistors and a 10.2-kΩ resistor. In comparison, if the inverting configuration were used with $R_1 = 1$ MΩ we would have required a feedback resistor of 100 MΩ, an impractically large value!

Before leaving this example it is insightful to inquire into the mechanism by which the circuit is able to realize a large voltage gain without using large resistances in the feedback path. Toward that end, observe that because of the virtual ground at the inverting input terminal of the op amp, R_2 and R_3 are in effect in parallel. Thus, by making R_3 lower than R_2 by, say, a factor k (i.e., where $k > 1$), R_3 is forced to carry a current k-times that in R_2. Thus, while $i_2 = i_1$, $i_3 = k i_1$ and $i_4 = (k+1)i_1$. It is the current multiplication by a factor of $(k+1)$ that enables a large voltage drop to develop across R_4 and hence a large v_O without using a large value for R_4. Notice also that the current through R_4 is independent of the value of R_4. It follows that the circuit can be used as a current amplifier as shown in Fig. 2.9.

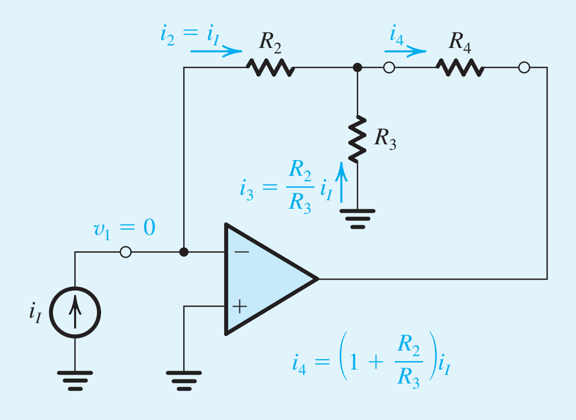


Figure 2.9 A current amplifier based on the circuit of Fig. 2.8. The amplifier delivers its output current to R_4. It has a current gain of $(1 + R_2/R_3)$, a zero input resistance, and an infinite output resistance. The load (R_4), however, must be floating (i.e., neither of its two terminals can be connected to ground).

EXERCISES

D2.4 Use the circuit of Fig. 2.5 to design an inverting amplifier having a gain of −10 and an input resistance of 100 kΩ. Give the values of R_1 and R_2.
Ans. $R_1 = 100$ kΩ; $R_2 = 1$ MΩ

2.5 The circuit shown in Fig. E2.5(a) can be used to implement a transresistance amplifier (see Table 1.1 in Section 1.5). Find the value of the input resistance R_i, the transresistance R_m, and the output resistance R_o of the transresistance amplifier. If the signal source shown in Fig. E2.5(b) is connected to the input of the transresistance amplifier, find the amplifier output voltage.
Ans. $R_i = 0$; $R_m = -10$ kΩ; $R_o = 0$; $v_O = -5$ V

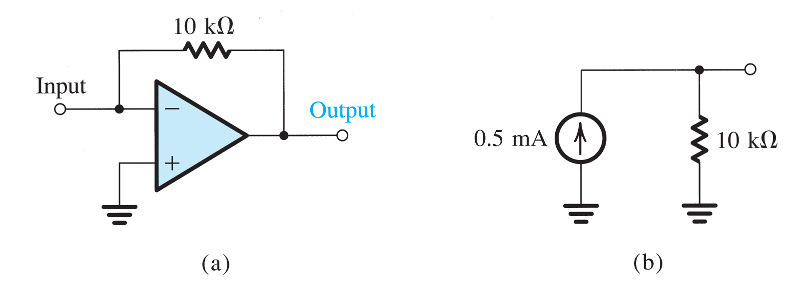


Figure E2.5

2.6 For the circuit in Fig. E2.6 determine the values of v_1, i_1, i_2, v_O, i_L, and i_O. Also determine the voltage gain v_O/v_I, current gain i_L/i_I, and power gain P_O/P_I.
Ans. 0 V; 1 mA; 1 mA; −10 V; −10 mA; −11 mA; −10 V/V (20 dB), −10 A/A (20 dB); 100 W/W (20 dB)

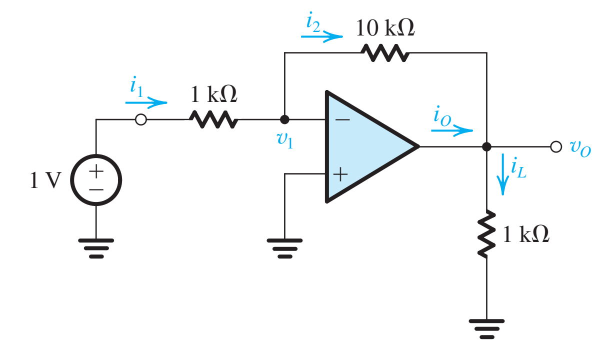


Figure E2.6

2.2.4 An Important Application—The Weighted Summer

A very important application of the inverting configuration is the weighted-summer circuit shown in Fig. 2.10. Here we have a resistance R_f in the negative-feedback path (as before), but we have a number of input signals $v_1, v_2, \ldots, v_n$ each applied to a corresponding resistor

$R_1, R_2, \ldots, R_n$, which are connected to the inverting terminal of the op amp. From our previous discussion, the ideal op amp will have a virtual ground appearing at its negative input terminal. Ohm's law then tells us that the currents $i_1, i_2, \ldots, i_n$ are given by

$$i_1 = \frac{v_1}{R_1}, \quad i_2 = \frac{v_2}{R_2}, \quad \ldots, \quad i_n = \frac{v_n}{R_n}$$

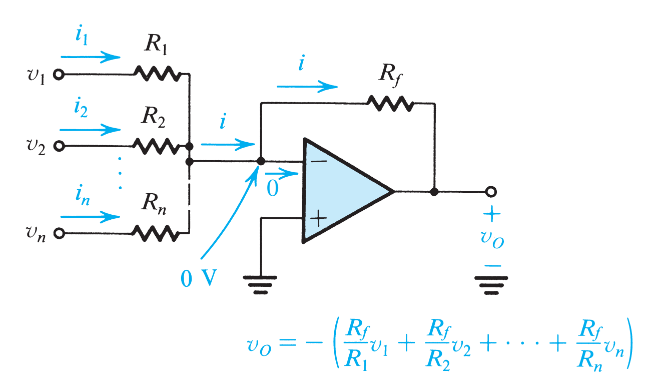


Figure 2.10 A weighted summer.

All these currents sum together to produce the current i,

$$i = i_1 + i_2 + \cdots + i_n \tag{2.6}$$

which will be forced to flow through R_f (since no current flows into the input terminals of an ideal op amp). The output voltage v_O may now be determined by another application of Ohm's law,

$$v_O = 0 - iR_f = -iR_f$$

Thus,

$$v_O = -\left(\frac{R_f}{R_1}v_1 + \frac{R_f}{R_2}v_2 + \cdots + \frac{R_f}{R_n}v_n\right) \tag{2.7}$$

That is, the output voltage is a weighted sum of the input signals $v_1, v_2, \ldots, v_n$. This circuit is therefore called a **weighted summer**. Note that each summing coefficient may be independently adjusted by adjusting the corresponding "feed-in" resistor (R_1 to R_n). This nice property, which greatly simplifies circuit adjustment, is a direct consequence of the virtual ground that exists at the inverting op-amp terminal. As the reader will soon come to appreciate, virtual grounds are extremely "handy." In the weighted summer of Fig. 2.10 all the summing coefficients must be of the same sign. The need occasionally arises for summing signals with opposite signs. Such a function can be implemented, however, using two op amps as shown in Fig. 2.11. Assuming ideal op amps, it can be easily shown that the output voltage is given by

$$v_O = v_1\left(\frac{R_a}{R_1}\right)\left(\frac{R_c}{R_b}\right) + v_2\left(\frac{R_a}{R_2}\right)\left(\frac{R_c}{R_b}\right) - v_3\left(\frac{R_c}{R_3}\right) - v_4\left(\frac{R_c}{R_4}\right) \tag{2.8}$$

Weighted summers are utilized in a variety of applications including in the design of audio systems, where they can be used in mixing signals originating from different musical instruments.

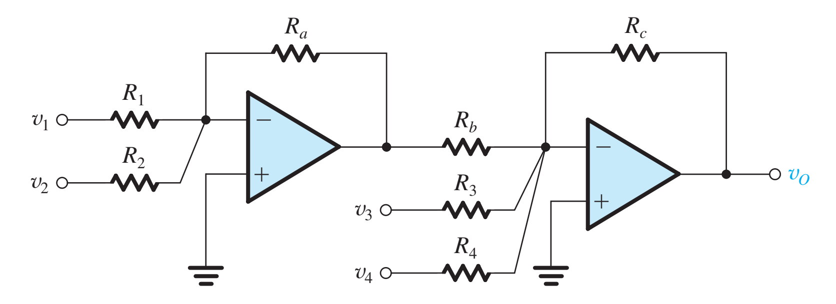


Figure 2.11 A weighted summer capable of implementing summing coefficients of both signs.

EXERCISES

D2.7 Design an inverting op-amp circuit to form the weighted sum v_O of two inputs v_1 and v_2. It is required that $v_O = -(v_1 + 5v_2)$. Choose values for R_1, R_2, and R_f so that for a maximum output voltage of 10 V the current in the feedback resistor will not exceed 1 mA.
Ans. A possible choice: $R_1 = 10\text{ k}\Omega, R_2 = 2\text{ k}\Omega$, and $R_f = 10\text{ k}\Omega$

D2.8 Use the idea presented in Fig. 2.11 to design a weighted summer that provides

$$v_O = 2v_1 + v_2 - 4v_3$$

Ans. A possible choice: $R_1 = 5\text{ k}\Omega$, $R_2 = 10\text{ k}\Omega$, $R_a = 10\text{ k}\Omega$, $R_b = 10\text{ k}\Omega$, $R_3 = 2.5\text{ k}\Omega$, $R_c = 10\text{ k}\Omega$

2.3 The Noninverting Configuration

The second closed-loop configuration we shall study is shown in Fig. 2.12. Here the input signal v_I is applied directly to the positive input terminal of the op amp while one terminal of R_1 is connected to ground.

2.3.1 The Closed-Loop Gain

Analysis of the noninverting circuit to determine its closed-loop gain (v_O/v_I) is illustrated in Fig. 2.13. Again the order of the steps in the analysis is indicated by circled numbers. Assuming that the op amp is ideal with infinite gain, a virtual short circuit exists between its two input terminals. Hence the difference input signal is

$$v_{Id} = \frac{v_O}{A} = 0 \qquad \text{for } A = \infty$$

Thus the voltage at the inverting input terminal will be equal to that at the noninverting input terminal, which is the applied voltage v_I. The current through R_1 can then be determined as v_I/R_1. Because of the infinite input impedance of the op amp, this current will flow through R_2, as shown in Fig. 2.13. Now the output voltage can be determined from

$$v_O = v_I + \left(\frac{v_I}{R_1}\right)R_2$$

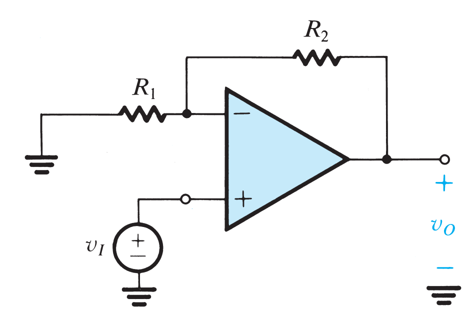


Figure 2.12 The noninverting configuration.

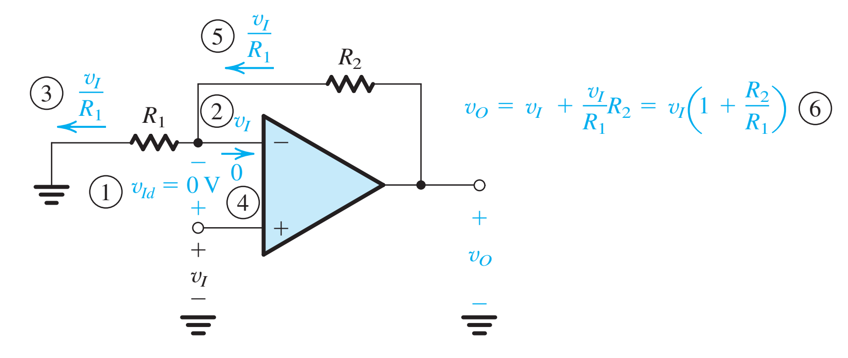


Figure 2.13 Analysis of the noninverting circuit. The sequence of the steps in the analysis is indicated by the circled numbers.

which yields



$$\frac{v_O}{v_I} = 1 + \frac{R_I}{R_2} \tag{2.9}$$

Further insight into the operation of the noninverting configuration can be obtained by considering the following: Since the current into the op-amp inverting input is zero, the circuit composed of R_1 and R_2 acts in effect as a voltage divider feeding a fraction of the output voltage back to the inverting input terminal of the op amp; that is,

$$v_1 = v_O\left(\frac{R_1}{R_1 + R_2}\right) \tag{2.10}$$

Then the infinite op-amp gain and the resulting virtual short circuit between the two input terminals of the op amp forces this voltage to be equal to that applied at the positive input terminal; thus,

$$v_O\left(\frac{R_1}{R_1 + R_2}\right) = v_I$$

which yields the gain expression given in Eq. (2.9).

This is an appropriate point to reflect further on the action of the negative feedback present in the noninverting circuit of Fig. 2.12. Let v_I increase. Such a change in v_I will cause v_{Id} to increase, and v_O will correspondingly increase as a result of the high (ideally infinite) gain of the op amp. However, a fraction of the increase in v_O will be fed back to the inverting input terminal of the op amp through the (R_1, R_2) voltage divider. The result of this feedback will be to counteract the increase in v_{Id}, driving v_{Id} back to zero, albeit at a higher value of v_O that corresponds to the increased value of v_I. This *degenerative* action of negative feedback gives it the alternative name **degenerative feedback**. Finally, note that the argument above applies equally well if v_I decreases. A formal and detailed study of feedback is presented in Chapter 11.

2.3.2 Effect of Finite Open-Loop Gain

As we have done for the inverting configuration, we now consider the effect of the finite op-amp open-loop gain A on the gain of the noninverting configuration. Assuming the op amp to be ideal except for having a finite open-loop gain A, it can be shown that the closed-loop gain of the noninverting amplifier circuit of Fig. 2.12 is given by

$$G \equiv \frac{v_O}{v_I} = \frac{1 + (R_2/R_1)}{1 + \dfrac{1 + (R_2/R_1)}{A}} \tag{2.11}$$

Observe that the denominator is identical to that for the case of the inverting configuration (Eq. 2.5). This is no coincidence; it is a result of the fact that both the inverting and the noninverting configurations have the same feedback loop, which can be readily seen if the input signal source is eliminated (i.e., short-circuited). The numerators, however, are different, for the numerator gives the ideal or nominal closed-loop gain ($-R_2/R_1$ for the inverting configuration, and $1 + R_2/R_1$ for the noninverting configuration). Finally, we note (with reassurance) that the gain expression in Eq. (2.11) reduces to the ideal value for $A = \infty$. In fact, it approximates the ideal value for

$$A \gg 1 + \frac{R_2}{R_1}$$

This is the same condition as in the inverting configuration, except that here the quantity on the right-hand side is the nominal closed-loop gain. The expressions for the actual and ideal values of the closed-loop gain G in Eqs. (2.11) and (2.9), respectively, can be used to determine the percentage error in G resulting from the finite op-amp gain A as

$$\text{Percent gain error } = -\frac{1 + (R_2/R_1)}{A + 1 + (R_2/R_1)} \times 100 \tag{2.12}$$

Thus, as an example, if an op amp with an open-loop gain of 1000 is used to design a noninverting amplifier with a nominal closed-loop gain of 10, we would expect the closed-loop gain to be about 1% below the nominal value.

2.3.3 Input and Output Resistance

The gain of the noninverting configuration is positive—hence the name *noninverting*. The input impedance of this closed-loop amplifier is ideally infinite, since no current flows into the positive input terminal of the op amp. The output of the noninverting amplifier is taken at the terminals of the ideal voltage source $A(v_2 - v_1)$ (see the op-amp equivalent circuit in Fig. 2.3), and thus the output resistance of the noninverting configuration is zero.

2.3.4 The Voltage Follower

The property of high input impedance is a very desirable feature of the noninverting configuration. It enables using this circuit as a buffer amplifier to connect a source with a high impedance to a low-impedance load. We discussed the need for buffer amplifiers in Section 1.5. In many applications the buffer amplifier is not required to provide any voltage gain; rather, it is used mainly as an impedance transformer or a power amplifier. In such cases we may make $R_2 = 0$ and $R_1 = \infty$ to obtain the **unity-gain amplifier** shown in Fig. 2.14(a). This circuit is commonly referred to as a **voltage follower**, since the output "follows" the input. In the ideal case, $v_O = v_I, R_{in} = \infty, R_{out} = 0$, and the follower has the equivalent circuit shown in Fig. 2.14(b).

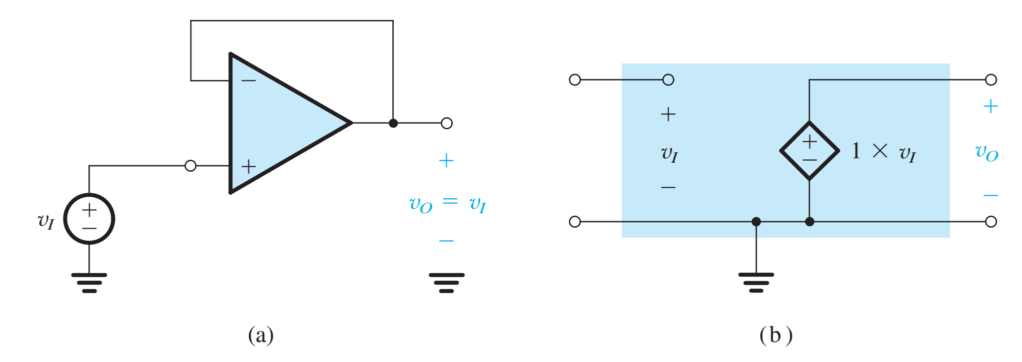


Figure 2.14 **(a)** The unity-gain buffer or follower amplifier. **(b)** Its equivalent circuit model.

Since in the voltage-follower circuit the entire output is fed back to the inverting input, the circuit is said to have 100% negative feedback. The infinite gain of the op amp then acts to make $v_{Id} = 0$ and hence $v_O = v_I$. Observe that the circuit is elegant in its simplicity!

Since the noninverting configuration has a gain greater than or equal to unity, depending on the choice of R_2/R_1, some prefer to call it "a follower with gain."

EXERCISES

2.9 Use the superposition principle to find the output voltage of the circuit shown in Fig. E2.9.
Ans. $v_O = 6v_1 + 4v_2$

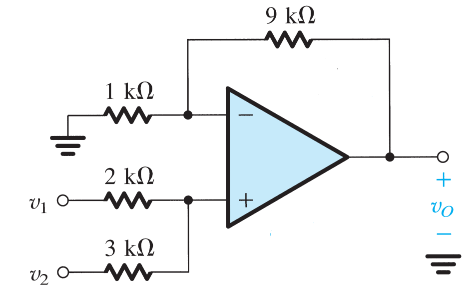


Figure E2.9

2.10 If in the circuit of Fig. E2.9 the 1-kΩ resistor is disconnected from ground and connected to a third signal source v_3, use superposition to determine v_O in terms of v_1, v_2, and v_3.
Ans. $v_O = 6v_1 + 4v_2 - 9v_3$

D2.11 Design a noninverting amplifier with a gain of 2. At the maximum output voltage of 10 V the current in the voltage divider is to be 10 μA.
Ans. $R_1 = R_2 = 0.5\ \text{M}\Omega$

2.12 (a) Show that if the op amp in the circuit of Fig. 2.12 has a finite open-loop gain A, then the closed-loop gain is given by Eq. (2.11). (b) For $R_1 = 1$ kΩ and $R_2 = 9$ kΩ find the percentage deviation ϵ of the closed-loop gain from the ideal value of $(1 + R_2/R_1)$ for the cases $A = 10^3, 10^4$, and 10^5. For $v_I = 1$ V, find in each case the voltage between the two input terminals of the op amp.
Ans. $\epsilon = -1\%, -0.1\%, -0.01\%$; $v_2 - v_1 = 9.9$ mV, 1 mV, 0.1 mV

2.13 For the circuit in Fig. E2.13 find the values of i_I, v_1, i_1, i_2, v_O, i_L, and i_O. Also find the voltage gain v_O/v_I, the current gain i_L/i_I, and the power gain P_L/P_I.
Ans. 0; 1 V; 1 mA; 1 mA; 10 V; 10 mA; 11 mA; 10 V/V (20 dB); ∞; ∞

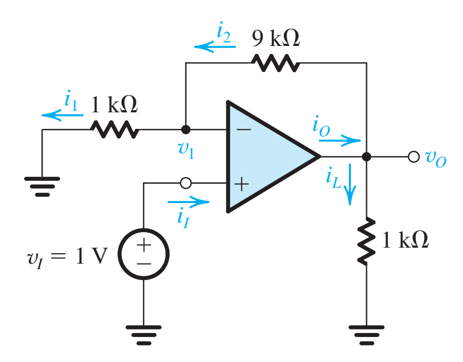


Figure E2.13

2.14 It is required to connect a transducer having an open-circuit voltage of 1 V and a source resistance of 1 MΩ to a load of 1-kΩ resistance. Find the load voltage if the connection is done (a) directly, and (b) through a unity-gain voltage follower.
Ans. (a) 1 mV; (b) 1 V

2.4 Difference Amplifiers

Having studied the two basic configurations of op-amp circuits together with some of their direct applications, we are now ready to consider a somewhat more involved but very important application. Specifically, we shall study the use of op amps to design difference or differential amplifiers.[2] A difference amplifier is one that responds to the difference between the two signals applied at its input and ideally rejects signals that are common to the two inputs. The representation of signals in terms of their differential and common-mode components was given in Fig. 2.4. It is repeated here in Fig. 2.15 with slightly different symbols to serve as the input signals for the difference amplifiers we are about to design. Although ideally the difference amplifier will amplify only the differential input signal v_{Id} and reject completely the common-mode input signal v_{Icm}, practical circuits will have an output voltage v_O given by

$$v_O = A_d v_{Id} + A_{cm} v_{Icm} \tag{2.13}$$

where A_d denotes the amplifier differential gain and A_{cm} denotes its common-mode gain (ideally zero). The efficacy of a differential amplifier is measured by the degree of its rejection of common-mode signals in preference to differential signals. This is usually quantified by a measure known as the **common-mode rejection ratio (CMRR)**, defined as

$$\text{CMRR} = 20 \log \frac{|A_d|}{|A_{cm}|} \tag{2.14}$$

[2]The terms *difference* and *differential* are usually used to describe somewhat different amplifier types. For our purposes at this point, the distinction is not sufficiently significant. We will be more precise near the end of this section.

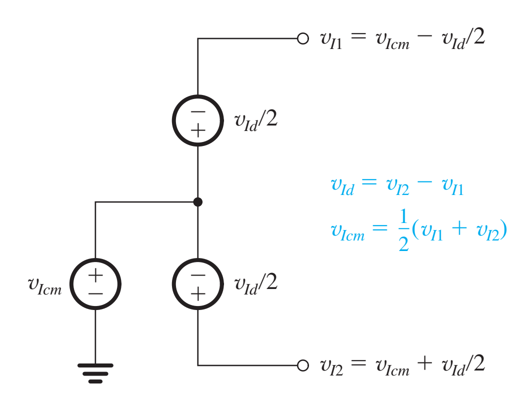


Figure 2.15 Representing the input signals to a differential amplifier in terms of their differential and common-mode components.

The need for difference amplifiers arises frequently in the design of electronic systems, especially those employed in instrumentation. As a common example, consider a transducer providing a small (e.g., 1 mV) signal between its two output terminals while each of the two wires leading from the transducer terminals to the measuring instrument may have a large interference signal (e.g., 1 V) relative to the circuit ground. The instrument front end obviously needs a difference amplifier.

Before we proceed any further we should address a question that the reader might have: The op amp is itself a difference amplifier; why not just use an op amp? The answer is that the very high (ideally infinite) gain of the op amp makes it impossible to use by itself. Rather, as we did before, we have to devise an appropriate feedback network to connect to the op amp to create a circuit whose closed-loop gain is finite, predictable, and stable.

2.4.1 A Single-Op-Amp Difference Amplifier

Our first attempt at designing a difference amplifier is motivated by the observation that the gain of the noninverting amplifier configuration is positive, $(1 + R_2/R_1)$, while that of the inverting configuration is negative, $(-R_2/R_1)$. Combining the two configurations together is then a step in the right direction—namely, getting the difference between two input signals. Of course, we have to make the two gain magnitudes equal in order to reject common-mode signals. This, however, can be easily achieved by attenuating the positive input signal to reduce the gain of the positive path from $(1+R_2/R_1)$ to (R_2/R_1). The resulting circuit would then look like that shown in Fig. 2.16, where the attenuation in the positive input path is achieved by the voltage divider (R_3, R_4). The proper ratio of this voltage divider can be determined from

$$\frac{R_4}{R_4+R_3}\left(1+\frac{R_2}{R_1}\right) = \frac{R_2}{R_1}$$

which can be put in the form

$$\frac{R_4}{R_4+R_3} = \frac{R_2}{R_2+R_1}$$

This condition is satisfied by selecting

$$\frac{R_4}{R_3} = \frac{R_2}{R_1} \tag{2.15}$$

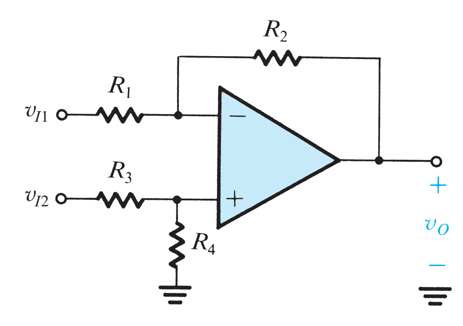


Figure 2.16 A difference amplifier.

This completes our work. However, we have perhaps proceeded a little too fast! Let's step back and verify that the circuit in Fig. 2.16 with R_3 and R_4 selected according to Eq. (2.15) does in fact function as a difference amplifier. Specifically, we wish to determine the output voltage v_O in terms of v_{I1} and v_{I2}. Toward that end, we observe that the circuit is linear, and thus we can use superposition.

To apply superposition, we first reduce v_{I2} to zero—that is, ground the terminal to which v_{I2} is applied—and then find the corresponding output voltage, which will be due entirely to v_{I1}. We denote this output voltage v_{O1}. Its value may be found from the circuit in Fig. 2.17(a), which we recognize as that of the inverting configuration. The existence of R_3 and R_4 does not affect the gain expression, since no current flows through either of them. Thus,

$$v_{O1} = -\frac{R_2}{R_1} v_{I1}$$

Next, we reduce v_{I1} to zero and evaluate the corresponding output voltage v_{O2}. The circuit will now take the form shown in Fig. 2.17(b), which we recognize as the noninverting configuration with an additional voltage divider, made up of R_3 and R_4, connected to the input v_{I2}. The output voltage v_{O2} is therefore given by

$$v_{O2} = v_{I2} \frac{R_4}{R_3 + R_4}\left(1 + \frac{R_2}{R_1}\right) = \frac{R_2}{R_1} v_{I2}$$

where we have utilized Eq. (2.15).

The superposition principle tells us that the output voltage v_O is equal to the sum of v_{O1} and v_{O2}. Thus we have

$$v_O = \frac{R_2}{R_1}(v_{I2} - v_{I1}) = \frac{R_2}{R_1} v_{Id} \tag{2.16}$$

Thus, as expected, the circuit acts as a difference amplifier with a differential gain A_d of

$$A_d = \frac{R_2}{R_1} \tag{2.17}$$

Of course this is predicated on the op amp being ideal and furthermore on the selection of R_3 and R_4 so that their ratio matches that of R_1 and R_2 (Eq. 2.15). To make this matching requirement a little easier to satisfy, we usually select

$$R_3 = R_1 \quad \text{and} \quad R_4 = R_2$$

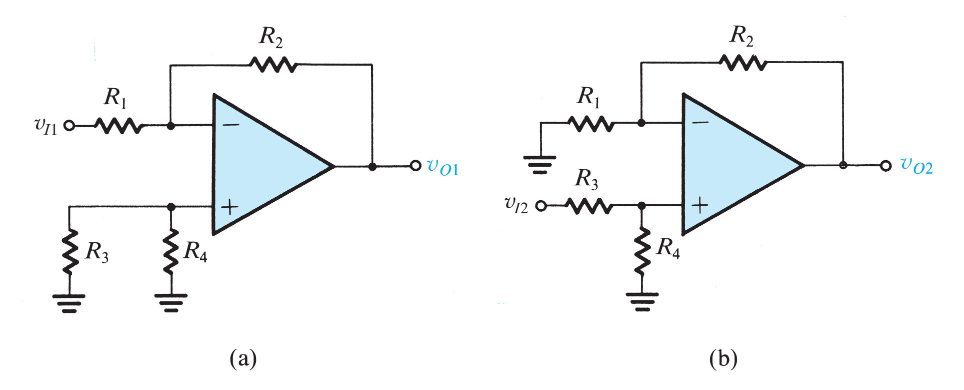


Figure 2.17 Application of superposition to the analysis of the circuit of Fig. 2.16.

Let's next consider the circuit with only a common-mode signal applied at the input, as shown in Fig. 2.18. The figure also shows some of the analysis steps. Thus,

$$i_1 = \frac{1}{R_1}\left[v_{Icm} - \frac{R_4}{R_4+R_3}v_{Icm}\right]$$

$$= v_{Icm}\frac{R_3}{R_4+R_3}\frac{1}{R_1} \tag{2.18}$$

The output voltage can now be found from

$$v_O = \frac{R_4}{R_4+R_3}v_{Icm} - i_2R_2$$

Substituting $i_2 = i_1$ and for i_1 from Eq. (2.18),

$$v_O = \frac{R_4}{R_4+R_3}v_{Icm} - \frac{R_2}{R_1}\frac{R_3}{R_4+R_3}v_{Icm}$$

$$= \frac{R_4}{R_4+R_3}\left(1 - \frac{R_2}{R_1}\frac{R_3}{R_4}\right)v_{Icm}$$

Thus,

$$A_{cm} \equiv \frac{v_O}{v_{Icm}} = \left(\frac{R_4}{R_4+R_3}\right)\left(1 - \frac{R_2}{R_1}\frac{R_3}{R_4}\right) \tag{2.19}$$

For the design with the resistor ratios selected according to Eq. (2.15), we obtain

$$A_{cm} = 0$$

as expected. Note, however, that any mismatch in the resistance ratios can make A_{cm} nonzero, and hence CMRR finite.

In addition to rejecting common-mode signals, a difference amplifier is usually required to have a high input resistance. To find the input resistance between the two input terminals

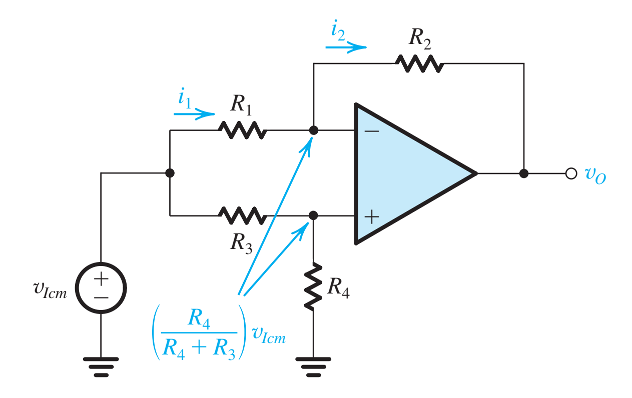


Figure 2.18 Analysis of the difference amplifier to determine its common-mode gain $A_{cm} \equiv v_O/v_{Icm}$.

(i.e., the resistance seen by v_{Id}), called the **differential input resistance** R_{id}, consider Fig. 2.19. Here we have assumed that the resistors are selected so that

$$R_3 = R_1 \quad \text{and} \quad R_4 = R_2$$

Now

$$R_{id} \equiv \frac{v_{Id}}{i_I}$$

Since the two input terminals of the op amp track each other in potential, we may write a loop equation and obtain

$$v_{Id} = R_1 i_I + 0 + R_1 i_I$$

Thus,

$$R_{id} = 2R_1 \tag{2.20}$$

Note that if the amplifier is required to have a large differential gain (R_2/R_1), then R_1 of necessity will be relatively small and the input resistance will be correspondingly low, a drawback of this circuit. Another drawback of the circuit is that it is not easy to vary the differential gain of the amplifier. Both of these drawbacks are overcome in the instrumentation amplifier discussed next.

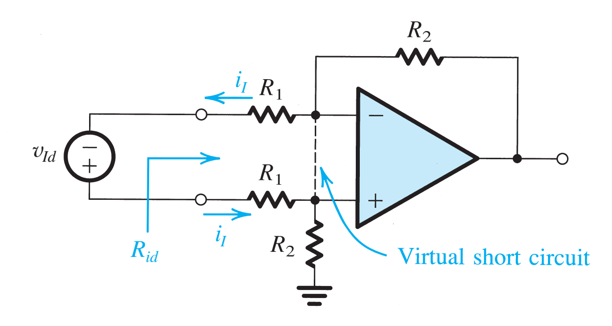


Figure 2.19 Finding the input resistance of the difference amplifier for the case $R_3 = R_1$ and $R_4 = R_2$.

EXERCISES

2.15 Consider the difference-amplifier circuit of Fig. 2.16 for the case $R_1 = R_3 = 2\ \text{k}\Omega$ and $R_2 = R_4 = 200\ \text{k}\Omega$. (a) Find the value of the differential gain A_d. (b) Find the value of the differential input resistance R_{id} and the output resistance R_o. (c) If the resistors have 1% tolerance (i.e., each can be within ±1% of its nominal value), use Eq. (2.19) to find the worst-case common-mode gain A_{cm} and hence the corresponding value of CMRR.
Ans. (a) 100 V/V (40 dB); (b) 4 kΩ, 0 Ω; (c) 0.04 V/V, 68 dB

D2.16 Find values for the resistances in the circuit of Fig. 2.16 so that the circuit behaves as a difference amplifier with an input resistance of 20 kΩ and a gain of 10.
Ans. $R_1 = R_3 = 10\ \text{k}\Omega$; $R_2 = R_4 = 100\ \text{k}\Omega$

2.4.2 A Superior Circuit—The Instrumentation Amplifier

The low-input-resistance problem of the difference amplifier of Fig. 2.16 can be solved by using voltage followers to buffer the two input terminals; that is, a voltage follower of the type in Fig. 2.14 is connected between each input terminal and the corresponding input terminal of the difference amplifier. However, if we are going to use two additional op amps, we should ask the question: Can we get more from them than just impedance buffering? An obvious answer would be that we should try to get some voltage gain. It is especially interesting that

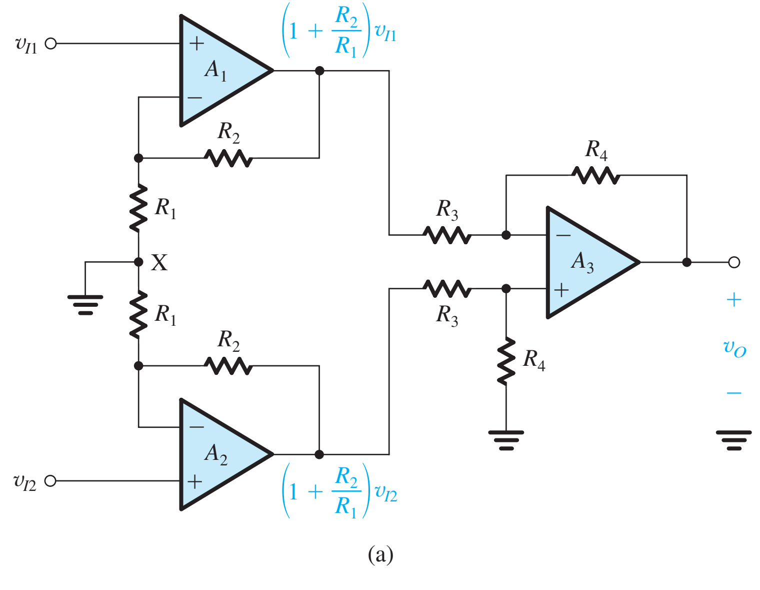


Figure 2.20 A popular circuit for an instrumentation amplifier. **(a)** Initial approach to the circuit. **(b)** The circuit in **(a)** with the connection between node X and ground removed and the two resistors R_1 and R_1 lumped together. This simple wiring change dramatically improves performance. **(c)** Analysis of the circuit in **(b)** assuming ideal op amps.

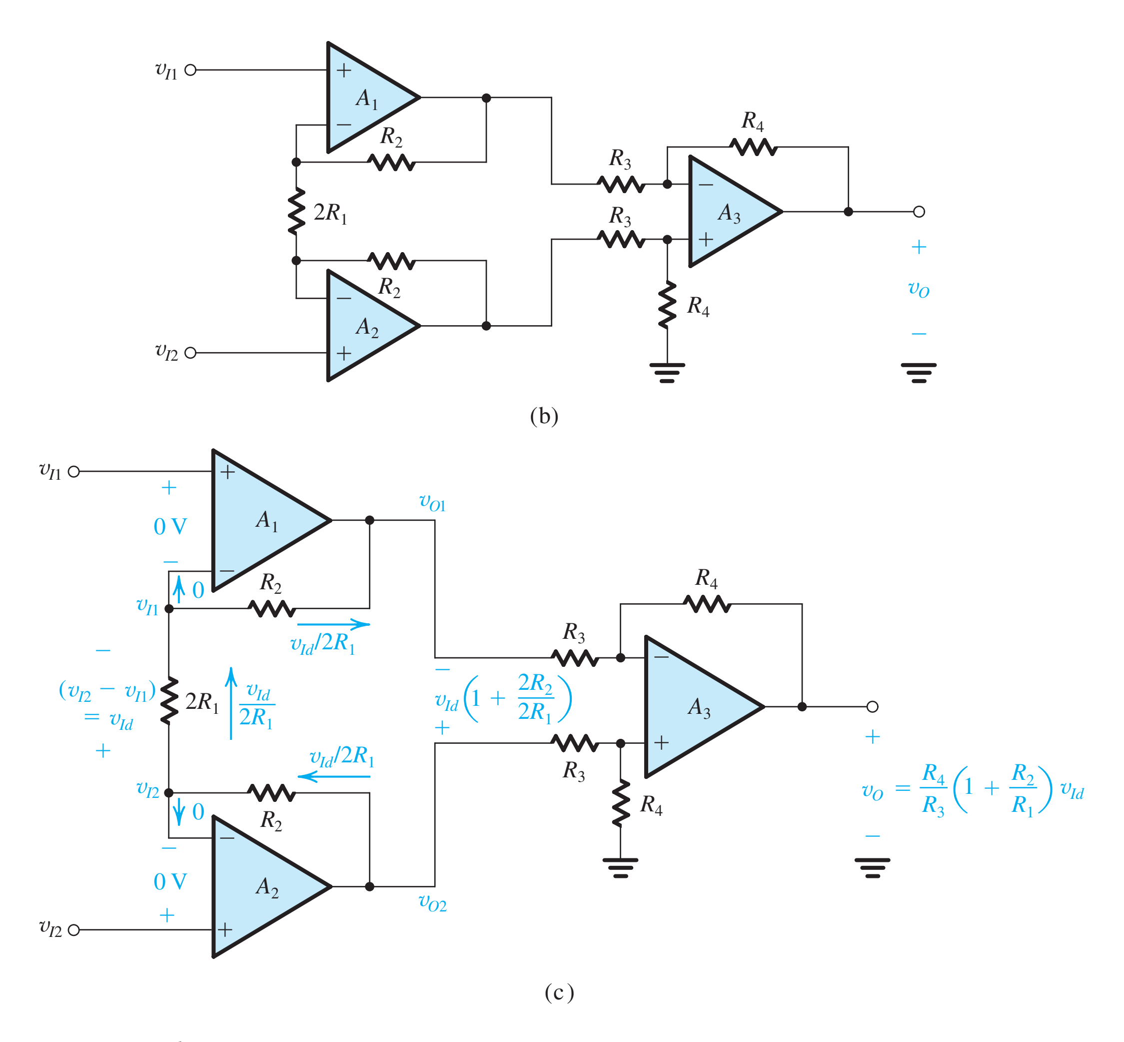


Figure 2.20 *continued*

we can achieve this without compromising the high input resistance simply by using followers with gain rather than unity-gain followers. Achieving some or indeed the bulk of the required gain in this new first stage of the differential amplifier eases the burden on the difference amplifier in the second stage, leaving it to its main task of implementing the differencing function and thus rejecting common-mode signals.

The resulting circuit is shown in Fig. 2.20(a). It consists of two stages in cascade. The first stage is formed by op amps A_1 and A_2 and their associated resistors, and the second stage is the by-now-familiar difference amplifier formed by op amp A_3 and its four associated resistors. Observe that as we set out to do, each of A_1 and A_2 is connected in the noninverting configuration and thus realizes a gain of $(1 + R_2/R_1)$. It follows that each of v_{I1} and v_{I2} is amplified by this factor, and the resulting amplified signals appear at the outputs of A_1 and A_2, respectively.

The difference amplifier in the second stage operates on the difference signal $(1 + R_2/R_1)(v_{I2} - v_{I1}) = (1 + R_2/R_1)v_{Id}$ and provides at its output

$$v_O = \frac{R_4}{R_3}\left(1 + \frac{R_2}{R_1}\right) v_{Id}$$

Thus the differential gain realized is

$$A_d = \left(\frac{R_4}{R_3}\right)\left(1 + \frac{R_2}{R_1}\right) \tag{2.21}$$

The common-mode gain will be zero because of the differencing action of the second-stage amplifier.

The circuit in Fig. 2.20(a) has the advantage of very high (ideally infinite) input resistance and high differential gain. Also, provided A_1 and A_2 and their corresponding resistors are matched, the two signal paths are symmetric—a definite advantage in the design of a differential amplifier. The circuit, however, has three major disadvantages:

1. The input common-mode signal v_{Icm} is amplified in the first stage by a gain equal to that experienced by the differential signal v_{Id}. This is a very serious issue, for it could result in the signals at the outputs of A_1 and A_3 being of such large magnitudes that the op amps saturate (more on op-amp saturation in Section 2.8). But even if the op amps do not saturate, the difference amplifier of the second stage will now have to deal with much larger common-mode signals, with the result that the CMRR of the overall amplifier will inevitably be reduced.
2. The two amplifier channels in the first stage have to be perfectly matched, otherwise a spurious signal may appear between their two outputs. Such a signal would get amplified by the difference amplifier in the second stage.
3. To vary the differential gain A_d, two resistors have to be varied simultaneously, say the two resistors labeled R_1. At each gain setting the two resistors have to be perfectly matched: a difficult task.

All three problems can be solved with a very simple wiring change: Simply disconnect the node between the two resistors labeled R_1, node X, from ground. The circuit with this small but functionally profound change is redrawn in Fig. 2.20(b), where we have lumped the two resistors (R_1 and R_1) together into a single resistor ($2R_1$).

Analysis of the circuit in Fig. 2.20(b), assuming ideal op amps, is straightforward, as is illustrated in Fig. 2.20(c). The key point is that the virtual short circuits at the inputs of op amps A_1 and A_2 cause the input voltages v_{I1} and v_{I2} to appear at the two terminals of resistor ($2R_1$). Thus the differential input voltage $v_{I2} - v_{I1} \equiv v_{Id}$ appears across $2R_1$ and causes a current $i = v_{Id}/2R_1$ to flow through $2R_1$ and the two resistors labeled R_2. This current in turn produces a voltage difference between the output terminals of A_1 and A_2 given by

$$v_{O2} - v_{O1} = \left(1 + \frac{2R_2}{2R_1}\right)v_{Id}$$

The difference amplifier formed by op amp A_3 and its associated resistors senses the voltage difference ($v_{O2} - v_{O1}$) and provides a proportional output voltage v_O:

$$v_O = \frac{R_4}{R_3}(v_{O2} - v_{O1})$$
$$= \frac{R_4}{R_3}\left(1 + \frac{R_2}{R_1}\right)v_{Id}$$

Thus the overall differential voltage-gain is given by

$$A_d \equiv \frac{v_O}{v_{Id}} = \frac{R_4}{R_3}\left(1+\frac{R_2}{R_1}\right) \tag{2.22}$$

Observe that proper differential operation does *not* depend on the matching of the two resistors labeled R_2. Indeed, if one of the two is of different value, say R_2', the expression for A_d becomes

$$A_d = \frac{R_4}{R_3}\left(1+\frac{R_2+R_2'}{2R_1}\right) \tag{2.23}$$

Consider next what happens when the two input terminals are connected together to a common-mode input voltage v_{Icm}. It is easy to see that an equal voltage appears at the negative input terminals of A_1 and A_2, causing the current through $2R_1$ to be zero. Thus there will be no current flowing in the R_2 resistors, and the voltages at the output terminals of A_1 and A_2 will be equal to the input (i.e., v_{Icm}). Thus the first stage no longer amplifies v_{Icm}; it simply propagates v_{Icm} to its two output terminals, where they are subtracted to produce a zero common-mode output by A_3. The difference amplifier in the second stage, however, now has a much improved situation at its input: The difference signal has been amplified by $(1+R_2/R_1)$ while the common-mode voltage remained unchanged.

Finally, we observe from the expression in Eq. (2.22) that the gain can be varied by changing only one resistor, $2R_1$. We conclude that this is an excellent differential amplifier circuit and is widely employed as an instrumentation amplifier, that is, as the input amplifier used in a variety of electronic instruments.

INTEGRATED INSTRUMENTATION AMPLIFIERS:

The conventional combination of three op amps and a number of precision resistors to form an instrumentation amplifier is an extremely powerful tool for the design of instruments for many applications. While the earliest applications used separate op amps and discrete resistors, fully integrated versions incorporating most required components in a single integrated-circuit package are increasingly available from many manufacturers. Low-power versions of these units are extremely important in the design of portable, wearable, and implantable medical monitoring devices, such as wristband activity monitors.

Example 2.3

Design the instrumentation amplifier circuit in Fig. 2.20(b) to provide a gain that can be varied over the range of 2 to 1000 utilizing a 100-kΩ variable resistance (a potentiometer, or "pot" for short).

Solution

It is usually preferable to obtain all the required gain in the first stage, leaving the second stage to perform the task of taking the difference between the outputs of the first stage and thereby rejecting the common-mode signal. In other words, the second stage is usually designed for a gain of 1. Adopting this approach, we select

Example 2.3 *continued*

all the second-stage resistors to be equal to a practically convenient value, say 10 kΩ. The problem then reduces to designing the first stage to realize a gain adjustable over the range of 2 to 1000. Implementing $2R_1$ as the series combination of a fixed resistor R_{1f} and the variable resistor R_{1v} obtained using the 100-kΩ pot (Fig. 2.21), we can write

$$1 + \frac{2R_2}{R_{1f} + R_{1v}} = 2 \text{ to } 1000$$

Thus,

$$1 + \frac{2R_2}{R_{1f}} = 1000$$

and

$$1 + \frac{2R_2}{R_{1f} + 100\text{ k}\Omega} = 2$$

These two equations yield $R_{1f} = 100.2\ \Omega$ and $R_2 = 50.050$ kΩ. Other practical values may be selected; for instance, $R_{1f} = 100\ \Omega$ and $R_2 = 49.9$ kΩ (both values are available as standard 1%-tolerance metal-film resistors; see Appendix J) results in a gain covering approximately the required range.

Figure 2.21 To make the gain of the circuit in Fig. 2.20(b) variable, $2R_1$ is implemented as the series combination of a fixed resistor R_{1f} and a variable resistor R_{1v}. Resistor R_{1f} ensures that the maximum available gain is limited.

EXERCISE

2.17 Consider the instrumentation amplifier of Fig. 2.20(b) with a common-mode input voltage of +5 V (dc) and a differential input signal of 10-mV-peak sine wave. Let $(2R_1) = 1$ kΩ, $R_2 = 0.5$ MΩ, and $R_3 = R_4 = 10$ kΩ. Find the voltage at every node in the circuit.
Ans. $v_{I1} = 5 - 0.005 \sin \omega t$; $v_{I2} = 5 + 0.005 \sin \omega t$; v_-(op amp A_1) $= 5 - 0.005 \sin \omega t$; v_-(op amp A_2) $= 5 + 0.005 \sin \omega t$; $v_{O1} = 5 - 5.005 \sin \omega t$; $v_{O2} = 5 + 5.005 \sin \omega t$; $v_-(A_3) = v_+(A_3) = 2.5 + 2.5025 \sin \omega t$; $v_O = 10.01 \sin \omega t$ (all in volts)

2.5 Integrators and Differentiators

The op-amp circuit applications we have studied thus far utilized resistors in the op-amp feedback path and in connecting the signal source to the circuit, that is, in the feed-in path. As a result, circuit operation has been (ideally) independent of frequency. By allowing the use of capacitors together with resistors in the feedback and feed-in paths of op-amp circuits, we open the door to a very wide range of useful and exciting applications of the op amp. We begin our study of op-amp–*RC* circuits by considering two basic applications, namely, signal integrators and differentiators.[3]

2.5.1 The Inverting Configuration with General Impedances

To begin with, consider the inverting closed-loop configuration with impedances $Z_1(s)$ and $Z_2(s)$ replacing resistors R_1 and R_2, respectively. The resulting circuit is shown in Fig. 2.22 and, for an ideal op amp, has the closed-loop gain or, more appropriately, the closed-loop transfer function

$$\frac{V_o(s)}{V_i(s)} = -\frac{Z_2(s)}{Z_1(s)} \tag{2.24}$$

As explained in Section 1.6, replacing s by $j\omega$ provides the transfer function for physical frequencies ω, that is, the transmission magnitude and phase for a sinusoidal input signal of frequency ω.

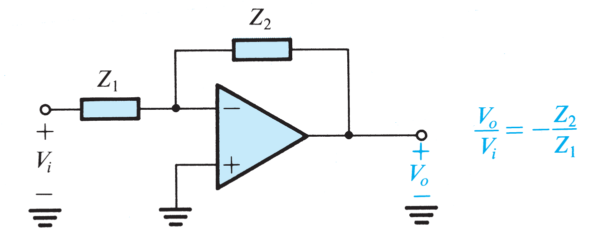


Figure 2.22 The inverting configuration with general impedances in the feedback and the feed-in paths.

[3] At this point, a review of Section 1.6 would be helpful. Also, an important fact to remember: Passing a constant current I through a capacitor C for a time t causes a change of It to accumulate on the capacitor. Thus the capacitor voltage changes by $\Delta V = \Delta Q/C = It/C$; that is, the capacitor voltage increases linearly with time.

EARLY OP AMPS AND ANALOG COMPUTATION:

In 1941, Karl D. Swartzel Jr. of Bell Labs patented "the summing amplifier," a high-gain dc inverting amplifier, intended to be used with negative feedback. This precursor of the op amp used three vacuum tubes (the predecessor of the transistor) and ±350-V power supplies to achieve a gain of 90 dB. Though lacking a differential input, it provided the usual applications of summation, integration, and general filtering using convenient passive resistive and capacitive components.

Soon after (1942), Loebe Julie, working with Professor John R. Regazzini at Columbia University, created a differential version, still using vacuum tubes. During World War II, these units were used extensively to provide analog computational functions in association with radar-directed antiaircraft firing control involving aircraft speed projection.

In the early 1950s, driven by the demonstrated wartime success of op-amp-based computation, general-purpose commercial systems called "analog computers" began to appear. They consisted of a few dozen op amps and associated passive components, including potentiometers; the interconnections required for programming were achieved with plug boards. These computers were used to solve differential equations.

Example 2.4

For the circuit in Fig. 2.23, derive an expression for the transfer function $V_o(s)/V_i(s)$. Show that the transfer function is that of a low-pass STC circuit. By expressing the transfer function in the standard form shown in Table 1.2 on page 36, find the dc gain and the 3-dB frequency. Design the circuit to obtain a dc gain of 40 dB, a 3-dB frequency of 1 kHz, and an input resistance of 1 kΩ. At what frequency does the magnitude of transmission become unity? What is the phase angle at this frequency?

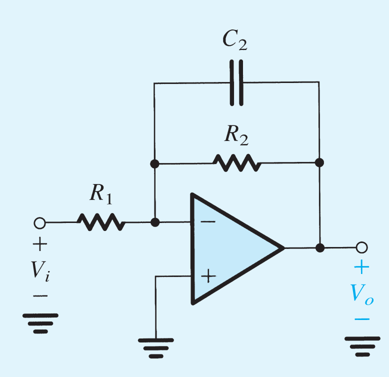


Figure 2.23 Circuit for Example 2.4.

Solution

To obtain the transfer function of the circuit in Fig. 2.23, we substitute in Eq. (2.24), $Z_1 = R_1$ and $Z_2 = R_2 \parallel (1/sC_2)$. Since Z_2 is the parallel connection of two components, it is more convenient to work in terms of Y_2; that is, we use the following alternative form of the transfer function:

$$\frac{V_o(s)}{V_i(s)} = -\frac{1}{Z_1(s)Y_2(s)}$$

and substitute $Z_1 = R_1$ and $Y_2(s) = (1/R_2) + sC_2$ to obtain

$$\frac{V_o(s)}{V_i(s)} = -\frac{1}{\dfrac{R_1}{R_2} + sC_2R_1}$$

This transfer function is of first order, has a finite dc gain (at $s = 0, V_o/V_i = -R_2/R_1$), and has zero gain at infinite frequency. Thus it is the transfer function of a low-pass STC network and can be expressed in the standard form of Table 1.2 as follows:

$$\frac{V_o(s)}{V_i(s)} = \frac{-R_2/R_1}{1 + sC_2R_2}$$

from which we find the dc gain K to be

$$K = -\frac{R_2}{R_1}$$

and the 3-dB frequency ω_0 as

$$\omega_0 = \frac{1}{C_2R_2}$$

We could have found all this from the circuit in Fig. 2.23 by inspection. Specifically, note that the capacitor behaves as an open circuit at dc; thus at dc the gain is simply $(-R_2/R_1)$. Furthermore, because there is a virtual ground at the inverting input terminal, the resistance seen by the capacitor is R_2, and thus the time constant of the STC network is C_2R_2.

Now to obtain a dc gain of 40 dB, that is, 100 V/V, we select $R_2/R_1 = 100$. For an input resistance of 1 kΩ, we select $R_1 = 1$ kΩ, and thus $R_2 = 100$ kΩ. Finally, for a 3-dB frequency $f_0 = 1$ kHz, we select C_2 from

$$2\pi \times 1 \times 10^3 = \frac{1}{C_2 \times 100 \times 10^3}$$

which yields $C_2 = 1.59$ nF.

The circuit has gain and phase Bode plots of the standard form in Fig. 1.23. As the gain falls off at the rate of –20 dB/decade, it will reach 0 dB in two decades, that is, at $f = 100f_0 = 100$ kHz. As Fig. 1.23(b) indicates, at such a frequency, which is much greater than f_0, the phase is approximately $-90°$. To this, however, we must add the 180° arising from the inverting nature of the amplifier (i.e., the negative sign in the transfer function expression). Thus at 100 kHz, the total phase shift will be $-270°$ or, equivalently, $+90°$.

2.5.2 The Inverting Integrator

By placing a capacitor in the feedback path (i.e., in place of Z_2 in Fig. 2.22) and a resistor at the input (in place of Z_1), we obtain the circuit of Fig. 2.24(a). We shall now show that this circuit realizes the mathematical operation of integration. Let the input be a time-varying function $v_I(t)$. The virtual ground at the inverting op-amp input causes $v_I(t)$ to appear in effect

across R, and thus the current $i_1(t)$ will be $v_I(t)/R$. This current flows through the capacitor C, causing charge to accumulate on C. If we assume that the circuit begins operation at time $t = 0$, then at an arbitrary time t the current $i_1(t)$ will have deposited on C a charge equal to $\int_0^t i_1(t)dt$. Thus the capacitor voltage $v_C(t)$ will change by $\frac{1}{C}\int_0^t i_1(t)dt$. If the initial voltage on C (at $t = 0$) is denoted V_C, then

$$v_C(t) = V_C + \frac{1}{C}\int_0^t i_1(t)dt$$

Now the output voltage $v_O(t) = -v_C(t)$; thus,

$$v_O(t) = -\frac{1}{CR}\int_0^t v_I(t)dt - V_C \tag{2.25}$$

Thus the circuit provides an output voltage that is proportional to the time integral of the input, with V_C being the initial condition of integration and CR the **integrator time constant**. Note that, as expected, there is a negative sign attached to the output voltage, and thus this integrator circuit is said to be an **inverting integrator**. It is also known as a **Miller integrator** after an early worker in this field.

The operation of the integrator circuit can be described alternatively in the frequency domain by substituting $Z_1(s) = R$ and $Z_2(s) = 1/sC$ in Eq. (2.24) to obtain the transfer function

$$\frac{V_o(s)}{V_i(s)} = -\frac{1}{sCR} \tag{2.26}$$

For physical frequencies, $s = j\omega$ and

$$\frac{V_o(j\omega)}{V_i(j\omega)} = -\frac{1}{j\omega CR} \tag{2.27}$$

Thus the integrator transfer function has magnitude

$$\left|\frac{V_o}{V_i}\right| = \frac{1}{\omega CR} \tag{2.28}$$

and phase

$$\phi = +90° \tag{2.29}$$

The Bode plot for the integrator magnitude response can be obtained by noting from Eq. (2.28) that as ω doubles (increases by an octave) the magnitude is halved (decreased by 6 dB). Thus the Bode plot is a straight line of slope –6 dB/octave (or, equivalently, –20 dB/ decade). This line (shown in Fig. 2.24b) intercepts the 0-dB line at the frequency that makes $|V_o/V_i| = 1$, which from Eq. (2.28) is

$$\omega_{\text{int}} = \frac{1}{CR} \tag{2.30}$$

The frequency ω_{int} is known as the **integrator frequency** and is simply the inverse of the integrator time constant.

Comparison of the frequency response of the integrator to that of an STC low-pass network indicates that the integrator behaves as a low-pass filter with a corner frequency of zero. Observe also that at $\omega = 0$, the magnitude of the integrator transfer function is infinite. This

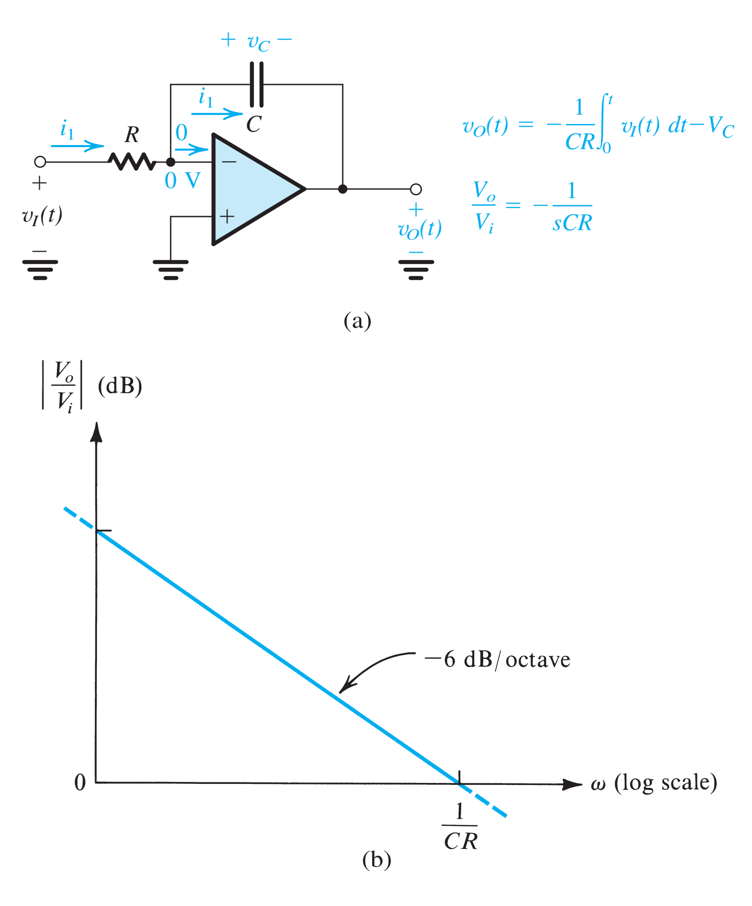


Figure 2.24 **(a)** The Miller or inverting integrator. **(b)** Frequency response of the integrator.

indicates that at dc the op amp is operating with an open loop. This should also be obvious from the integrator circuit itself. Reference to Fig. 2.24(a) shows that the feedback element is a capacitor, and thus at dc, where the capacitor behaves as an open circuit, there is no negative feedback! This is a very significant observation and one that indicates a source of problems with the integrator circuit: Any tiny dc component in the input signal will theoretically produce an infinite output. Of course, no infinite output voltage results in practice; rather, the output of the amplifier saturates at a voltage close to the op-amp positive or negative power supply (L_+ or L_-), depending on the polarity of the input dc signal.

The dc problem of the integrator circuit can be alleviated by connecting a resistor R_F across the integrator capacitor C, as shown in Fig. 2.25, and thus the gain at dc will be $-R_F/R$ rather than infinite. Such a resistor provides a dc feedback path. Unfortunately, however, the integration is no longer ideal, and the lower the value of R_F, the less ideal the integrator circuit becomes. This is because R_F causes the frequency of the integrator pole to move from its ideal location at $\omega = 0$ to one determined by the corner frequency of the STC network (R_F, C). Specifically, the integrator transfer function becomes

$$\frac{V_o(s)}{V_i(s)} = -\frac{R_F/R}{1 + sCR_F}$$

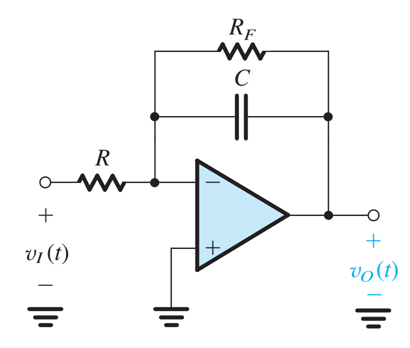


Figure 2.25 The Miller integrator with a large resistance R_F connected in parallel with C in order to provide negative feedback and hence finite gain at dc.

as opposed to the ideal function of $-1/sCR$. The lower the value we select for R_F, the higher the corner frequency ($1/CR_F$) will be and the more nonideal the integrator becomes. Thus selecting a value for R_F presents the designer with a trade-off between dc performance and signal performance. The effect of R_F on integrator performance is investigated further in Example 2.5.

Example 2.5

Find the output produced by a Miller integrator in response to an input pulse of 1-V height and 1-ms width [Fig. 2.26(a)]. Let $R = 10$ kΩ and $C = 10$ nF. If the integrator capacitor is shunted by a 1-MΩ resistor, how will the response be modified? The op amp is specified to saturate at ±13 V.

Solution

In response to a 1-V, 1-ms input pulse, the integrator output will be

$$v_O(t) = -\frac{1}{CR}\int_0^t 1\,dt, \qquad 0 \le t \le 1 \text{ ms}$$

where we have assumed that the initial voltage on the integrator capacitor is 0. For $C = 10$ nF and $R = 10$ kΩ, $CR = 0.1$ ms, and

$$v_O(t) = -10t, \qquad 0 \le t \le 1 \text{ ms}$$

which is the linear ramp shown in Fig. 2.26(b). It reaches a magnitude of −10 V at $t = 1$ ms and remains constant thereafter.

That the output is a linear ramp should also be obvious from the fact that the 1-V input pulse produces a constant current through the capacitor of 1 V/10 kΩ = 0.1 mA. This constant current $I = 0.1$ mA supplies the capacitor with a charge It, and thus the capacitor voltage changes linearly as (It/C), resulting in $v_O = -(I/C)t$. It is worth remembering that charging a capacitor with a constant current produces a linear voltage across it.

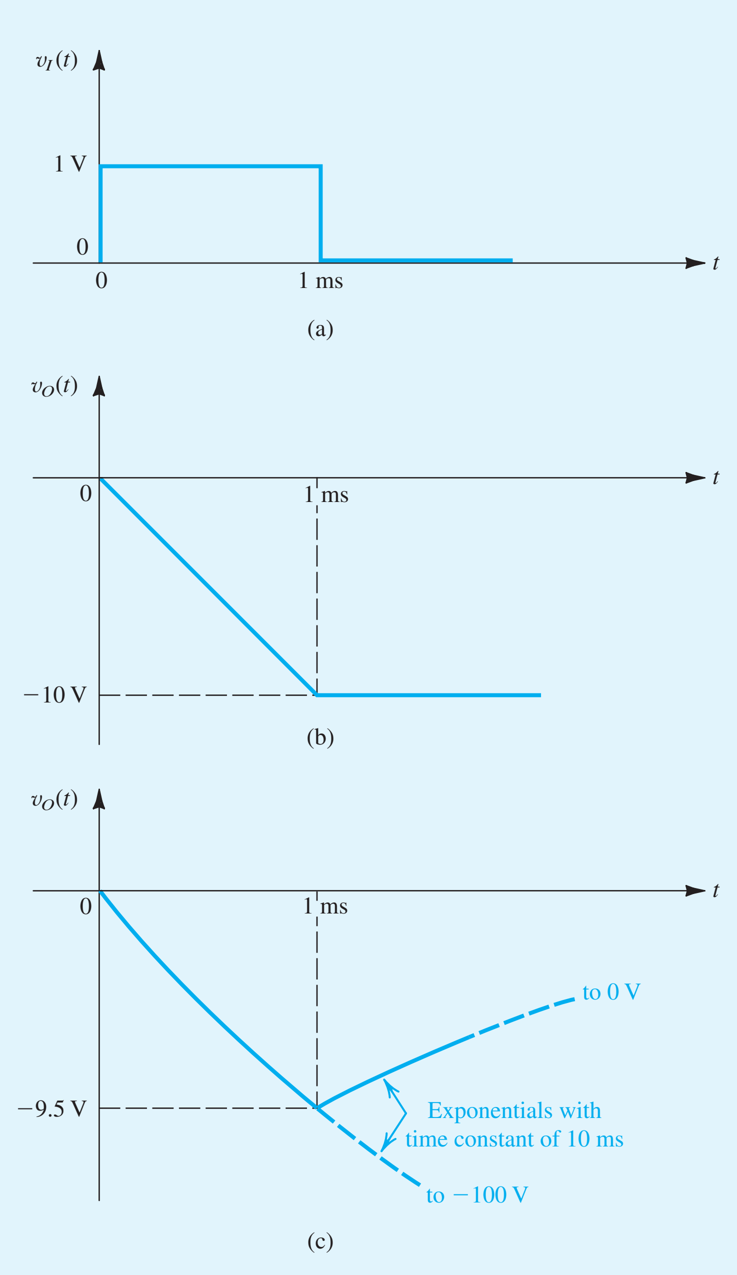


Figure 2.26 Waveforms for Example 2.5: **(a)** Input pulse. **(b)** Output linear ramp of ideal integrator with time constant of 0.1 ms. **(c)** Output exponential ramp with resistor R_F connected across integrator capacitor.

Next consider the situation with resistor $R_F = 1\ \text{M}\Omega$ connected across C. As before, the 1-V pulse will provide a constant current $I = 0.1$ mA. Now, however, this current is supplied to an STC network composed of R_F in parallel with C. Thus, the output will be an exponential heading toward -100 V with

Example 2.5 *continued*

a time constant of $CR_F = 10 \times 10^{-9} \times 1 \times 10^6 = 10$ ms,

$$v_O(t) = -100(1 - e^{-t/10}), \qquad 0 \leq t \leq 1 \text{ ms}$$

Of course, the exponential will be interrupted at the end of the pulse, that is, at $t = 1$ ms, and the output will reach the value

$$v_O(1 \text{ ms}) = -100(1 - e^{-1/10}) = -9.5 \text{ V}$$

The output waveform is shown in Fig. 2.26(c), from which we see that including R_F causes the ramp to be slightly rounded such that the output reaches only −9.5 V, 0.5 V short of the ideal value of −10 V. Furthermore, for $t > 1$ ms, the capacitor discharges through R_F with the relatively long time constant of 10 ms. Finally, we note that op-amp saturation, specified to occur at ±13 V, has no effect on the operation of this circuit.

The preceding example hints at an important application of integrators, namely, their use in providing triangular waveforms in response to square-wave inputs. This application is explored in Exercise 2.18. Integrators have many other applications, including their use in the design of filters (Chapter 17).

2.5.3 The Op-Amp Differentiator

Interchanging the location of the capacitor and the resistor of the integrator circuit results in the circuit in Fig. 2.27(a), which performs the mathematical function of differentiation. To see how this comes about, let the input be the time-varying function $v_I(t)$, and note that the virtual ground at the inverting input terminal of the op amp causes $v_I(t)$ to appear in effect across the capacitor C. Thus the current through C will be $C(dv_I/dt)$, and this current flows through the feedback resistor R providing at the op-amp output a voltage $v_O(t)$,

$$v_O(t) = -CR\frac{dv_I(t)}{dt} \tag{2.31}$$

The frequency-domain transfer function of the differentiator circuit can be found by substituting in Eq. (2.24), $Z_1(s) = 1/sC$ and $Z_2(s) = R$ to obtain

$$\frac{V_o(s)}{V_i(s)} = -sCR \tag{2.32}$$

which for physical frequencies $s = j\omega$ yields

$$\frac{V_o(j\omega)}{V_i(j\omega)} = -j\omega CR \tag{2.33}$$

Thus the transfer function has magnitude

$$\left|\frac{V_o}{V_i}\right| = \omega CR \tag{2.34}$$

and phase

$$\phi = -90° \quad (2.35)$$

The Bode plot of the magnitude response can be found from Eq. (2.34) by noting that for an octave increase in ω, the magnitude doubles (increases by 6 dB). Thus the plot is simply a straight line of slope +6 dB/octave (or, equivalently, +20 dB/decade) intersecting the 0-dB line (where $|V_o/V_i| = 1$) at $\omega = 1/CR$, where CR is the **differentiator time constant** [see Fig. 2.27(b)].

The frequency response of the differentiator can be thought of as the response of an STC high-pass filter with a corner frequency at infinity (refer to Fig. 1.24). Finally, we should note that the very nature of a differentiator circuit causes it to be a "noise magnifier." This is due to the spike introduced at the output every time there is a sharp change in $v_I(t)$; such a change could be interference coupled electromagnetically ("picked up") from adjacent signal sources. For this reason and because they suffer from stability problems (Chapter 11), differentiator circuits are generally avoided in practice. When the circuit of Fig. 2.27(a) is used, it is usually necessary to connect a small-valued resistor in series with the capacitor. This modification, unfortunately, turns the circuit into a nonideal differentiator.

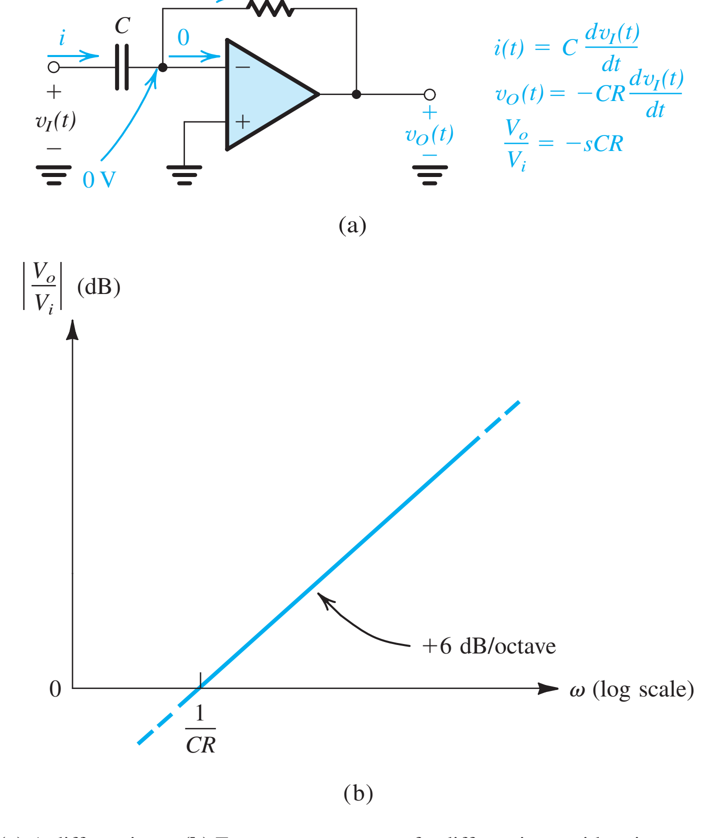


Figure 2.27 **(a)** A differentiator. **(b)** Frequency response of a differentiator with a time constant CR.

EXERCISES

2.18 Consider a symmetrical square wave of 20-V peak-to-peak, 0 average, and 2-ms period applied to a Miller integrator. Find the value of the time constant *CR* such that the triangular waveform at the output has a 20-V peak-to-peak amplitude.
Ans. 0.5 ms

D2.19 Use an ideal op amp to design an inverting integrator with an input resistance of 10 kΩ and an integration time constant of 10^{-3} s. What is the gain magnitude and phase angle of this circuit at 10 rad/s and at 1 rad/s? What is the frequency at which the gain magnitude is unity?
Ans. $R = 10$ kΩ, $C = 0.1$ μF; at $\omega = 10$ rad/s: $|V_o/V_i| = 100$ V/V and $\phi = +90°$; at $\omega = 1$ rad/s: $|V_o/V_i| = 1000$ V/V and $\phi = +90°$; 1000 rad/s

D2.20 Design a differentiator to have a time constant of 10^{-2} s and an input capacitance of 0.01 μF. What is the gain magnitude and phase of this circuit at 10 rad/s, and at 10^3 rad/s? In order to limit the high-frequency gain of the differentiator circuit to 100, a resistor is added in series with the capacitor. Find the required resistor value.
Ans. $C = 0.01$ μF; $R = 1$ MΩ; at $\omega = 10$ rad/s: $|V_o/V_i| = 0.1$ V/V and $\phi = -90°$; at $\omega = 1000$ rad/s: $|V_o/V_i| = 10$ V/V and $\phi = -90°$; 10 kΩ

2.6 DC Imperfections

Thus far we have considered the op amp to be ideal. The only exception has been a brief discussion of the effect of the op-amp finite gain *A* on the closed-loop gain of the inverting and noninverting configurations. Although in many applications the assumption of an ideal op amp is not a bad one, a circuit designer has to be thoroughly familiar with the characteristics of practical op amps and the effects of such characteristics on the performance of op-amp circuits. Only then will the designer be able to use the op amp intelligently, especially if the application at hand is not a straightforward one. The nonideal properties of op amps will, of course, limit the range of operation of the circuits analyzed in the previous examples.

In this and the two sections that follow, we consider some of the important nonideal properties of the op amp.[4] We do this by treating one nonideality at a time, beginning in this section with the dc problems to which op amps are susceptible.

2.6.1 Offset Voltage

Because op amps are direct-coupled devices with large gains at dc, they are prone to dc problems. The first such problem is the dc offset voltage. To understand this problem consider

[4]We should note that real op amps have nonideal effects additional to those discussed in this chapter. These include finite (nonzero) common-mode gain or, equivalently, noninfinite CMRR, noninfinite input resistance, and nonzero output resistance. The effect of these, however, on the performance of most of the closed-loop circuits studied here is not very significant, and their study will be postponed to later chapters (in particular, Chapters 9, 10, and 13).

the following *conceptual* experiment: If the two input terminals of the op amp are tied together and connected to ground, it will be found that despite the fact that $v_{Id} = 0$, a finite dc voltage exists at the output. In fact, if the op amp has a high dc gain, the output will be at either the positive or negative saturation level. The op-amp output can be brought back to its ideal value of 0 V by connecting a dc voltage source of appropriate polarity and magnitude between the two input terminals of the op amp. This external source balances out the input offset voltage of the op amp. It follows that the **input offset voltage** (V_{OS}) must be of equal magnitude and of opposite polarity to the voltage we applied externally.

The input offset voltage arises as a result of the unavoidable mismatches present in the input differential stage inside the op amp. In later chapters (in particular Chapters 9 and 13) we shall study this topic in detail. Here, however, our concern is to investigate the effect of V_{OS} on the operation of closed-loop op-amp circuits. Toward that end, we note that general-purpose op amps exhibit V_{OS} in the range of 1 mV to 5 mV. Also, the value of V_{OS} depends on temperature. The op-amp data sheets usually specify typical and maximum values for V_{OS} at room temperature as well as the temperature coefficient of V_{OS} (usually in μV/°C). They do not, however, specify the polarity of V_{OS} because the component mismatches that give rise to V_{OS} are obviously not known a priori; different units of the same op-amp type may exhibit either a positive or a negative V_{OS}.

To analyze the effect of V_{OS} on the operation of op-amp circuits, we need a circuit model for the op amp with input offset voltage. Such a model is shown in Fig. 2.28. It consists of a dc source of value V_{OS} placed in series with the positive input lead of an offset-free op amp. The justification for this model follows from the description above.

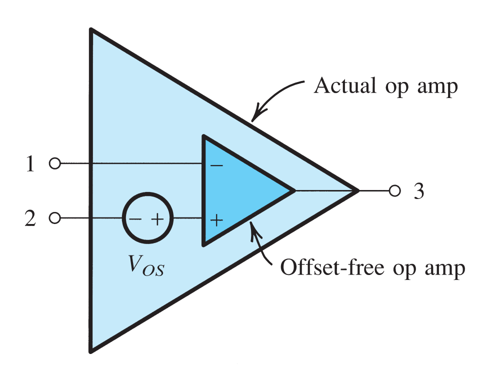


Figure 2.28 Circuit model for an op amp with input offset voltage V_{OS}.

EXERCISE

2.21 Use the model of Fig. 2.28 to sketch the transfer characteristic v_O versus v_{Id} ($v_O \equiv v_3$ and $v_{Id} \equiv v_2 - v_1$) of an op amp having an open-loop dc gain $A_0 = 10^4$ V/V, output saturation levels of ±10 V, and V_{OS} of +5 mV.

Ans. See Fig. E2.21. Observe that true to its name, the input offset voltage causes an offset in the voltage-transfer characteristic; rather than passing through the origin it is now shifted to the left by V_{OS}.

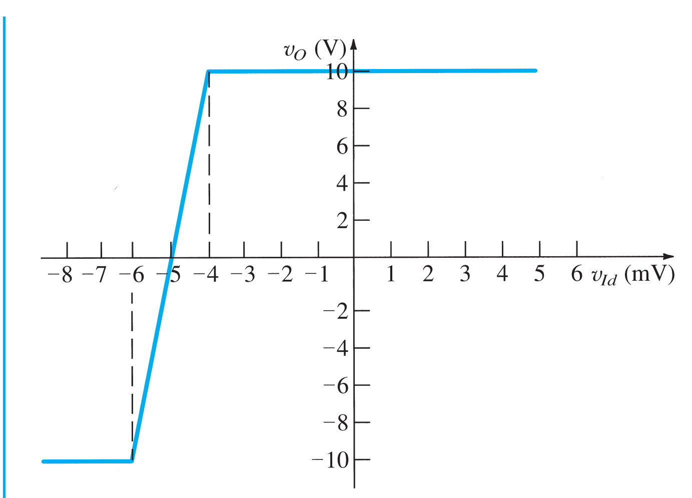


Figure E2.21 Transfer characteristic of an op amp with $V_{OS} = 5$ mV.

Analysis of op-amp circuits to determine the effect of the op-amp V_{OS} on their performance is straightforward: The input voltage signal source is short-circuited and the op amp is replaced with the model of Fig. 2.28. (Eliminating the input signal, done to simplify matters, is based on the principle of superposition.) Following this procedure, we find that both the inverting and the noninverting amplifier configurations result in the same circuit, that shown in Fig. 2.29, from which the output dc voltage due to V_{OS} is found to be

$$V_O = V_{OS}\left[1 + \frac{R_2}{R_1}\right] \tag{2.36}$$

This output dc voltage can have a large magnitude. For instance, a noninverting amplifier with a closed-loop gain of 1000, when constructed from an op amp with a 5-mV input offset voltage, will have a dc output voltage of +5 V or −5 V (depending on the polarity of V_{OS}) rather than the ideal value of 0 V. Now, when an input signal is applied to the amplifier, the corresponding signal output will be superimposed on the 5-V dc. Obviously then, the

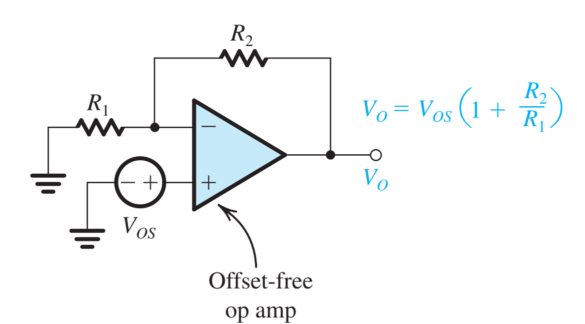


Figure 2.29 Evaluating the output dc offset voltage due to V_{OS} in a closed-loop amplifier.

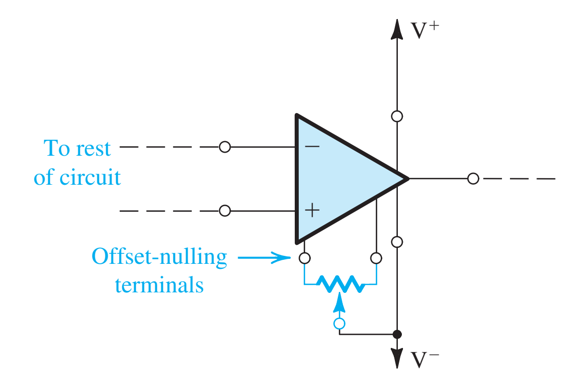


Figure 2.30 The output dc offset voltage of an op amp can be trimmed to zero by connecting a potentiometer to the two offset-nulling terminals. The wiper of the potentiometer is connected to the negative supply of the op amp.

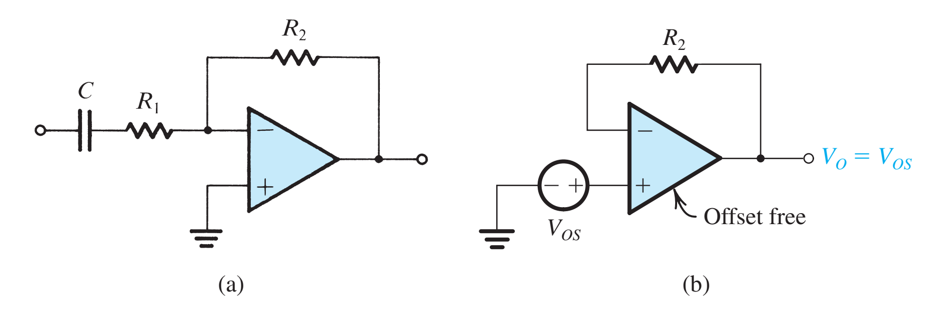


Figure 2.31 **(a)** A capacitively coupled inverting amplifier. **(b)** The equivalent circuit for determining its dc output offset voltage V_O.

allowable signal swing at the output will be reduced. Even worse, if the signal to be amplified is dc, we would not know whether the output is due to V_{OS} or to the signal!

Some op amps are provided with two additional terminals to which a specified circuit can be connected to trim to zero the output dc voltage due to V_{OS}. Figure 2.30 shows such an arrangement that is typically used with general-purpose op amps. A potentiometer is connected between the offset-nulling terminals with the wiper of the potentiometer connected to the op-amp negative supply. Moving the potentiometer wiper introduces an imbalance that counteracts the asymmetry present in the internal op-amp circuitry and that gives rise to V_{OS}. We shall return to this point in the context of our study of the internal circuitry of op amps in Chapter 13. It should be noted, however, that even though the dc output offset can be trimmed to zero, the problem remains of the variation (or drift) of V_{OS} with temperature.

One way to overcome the dc offset problem is by capacitively coupling the amplifier. This, however, will be possible only in applications where the closed-loop amplifier is not required to amplify dc or very-low-frequency signals. Figure 2.31(a) shows a capacitively coupled amplifier. Because of its infinite impedance at dc, the coupling capacitor will cause the gain to be zero at dc. As a result, the equivalent circuit for determining the dc output voltage resulting from the op-amp input offset voltage V_{OS} will be that shown in Fig. 2.31(b). Thus V_{OS} sees in effect a unity-gain voltage follower, and the dc output voltage V_O will be equal to V_{OS} rather than $V_{OS}(1+R_2/R_1)$, which is the case without the coupling capacitor. As far as input signals are concerned, the coupling capacitor C forms together with R_1 an STC high-pass circuit with a corner frequency of $\omega_0 = 1/CR_1$. Thus the gain of the capacitively

coupled amplifier will fall off at the low-frequency end [from a magnitude of $(1+R_2/R_1)$ at high frequencies] and will be 3 dB down at ω_0.

EXERCISES

2.22 Consider an inverting amplifier with a nominal gain of 1000 constructed from an op amp with an input offset voltage of 3 mV and with output saturation levels of ±10 V. (a) What is (approximately) the peak sine-wave input signal that can be applied without output clipping? (b) If the effect of V_{OS} is nulled at room temperature (25°C), how large an input can one now apply if: (i) the circuit is to operate at a constant temperature? (ii) the circuit is to operate at a temperature in the range 0°C to 75°C and the temperature coefficient of V_{OS} is 10 μV/°C?
Ans. (a) 7 mV; (b) 10 mV, 9.5 mV

2.23 Consider the same amplifier as in Exercise 2.22—that is, an inverting amplifier with a nominal gain of 1000 constructed from an op amp with an input offset voltage of 3 mV and with output saturation levels of ±10 V—except here let the amplifier be capacitively coupled as in Fig. 2.31(a). (a) What is the dc offset voltage at the output, and what (approximately) is the peak sine-wave signal that can be applied at the input without output clipping? Is there a need for offset trimming? (b) If $R_1 = 1$ kΩ and $R_2 = 1$ MΩ, find the value of the coupling capacitor C_1 that will ensure that the gain will be greater than 57 dB down to 100 Hz.
Ans. (a) 3 mV, 10 mV, no need for offset trimming; (b) 1.6 μF

2.6.2 Input Bias and Offset Currents

The second dc problem encountered in op amps is illustrated in Fig. 2.32. In order for the op amp to operate, its two input terminals have to be supplied with dc currents, termed the **input bias currents**.[5] In Fig. 2.32 these two currents are represented by two current sources, I_{B1} and I_{B2}, connected to the two input terminals. It should be emphasized that the input bias currents are independent of the fact that a real op amp has finite (though large) input resistance (not shown in Fig. 2.32). The op-amp manufacturer usually specifies the average value of I_{B1} and I_{B2} as well as their expected difference. The average value I_B is called the **input bias current**,

$$I_B = \frac{I_{B1}+I_{B2}}{2}$$

and the difference is called the **input offset current** and is given by

$$I_{OS} = |I_{B1} - I_{B2}|$$

Typical values for general-purpose op amps that use bipolar transistors are $I_B = 100$ nA and $I_{OS} = 10$ nA.

[5]This is the case for op amps constructed using bipolar junction transistors (BJTs). Those using MOSFETs in the first (input) stage do not draw an appreciable input bias current; nevertheless, the input terminals should have continuous dc paths to ground. More on this in later chapters.

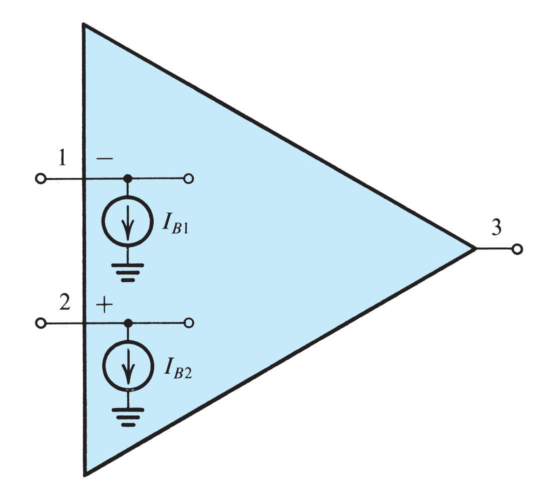


Figure 2.32 The op-amp input bias currents represented by two current sources I_{B1} and I_{B2}.

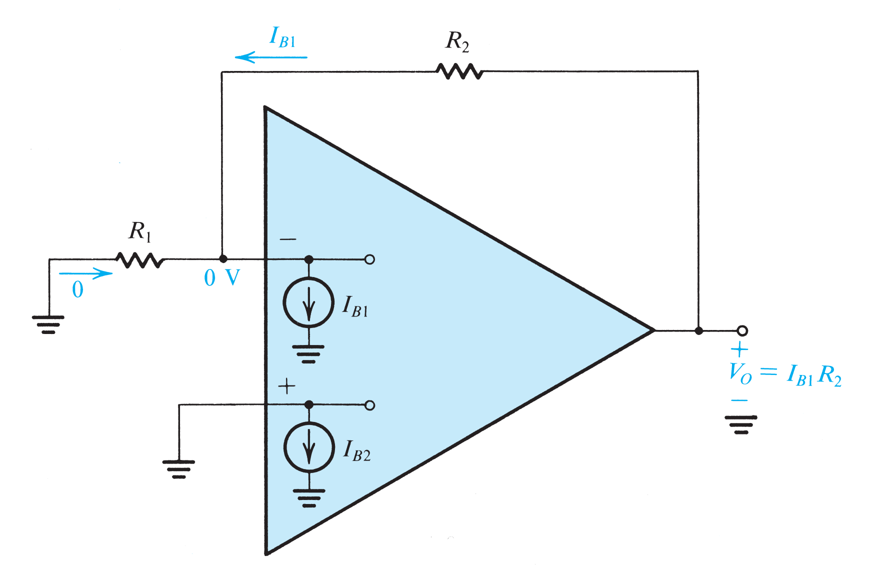


Figure 2.33 Analysis of the closed-loop amplifier, taking into account the input bias currents.

We now wish to find the dc output voltage of the closed-loop amplifier due to the input bias currents. To do this we ground the signal source and obtain the circuit shown in Fig. 2.33 for both the inverting and noninverting configurations. As shown in Fig. 2.33, the output dc voltage is given by

$$V_O = I_{B1}R_2 \simeq I_B R_2 \tag{2.37}$$

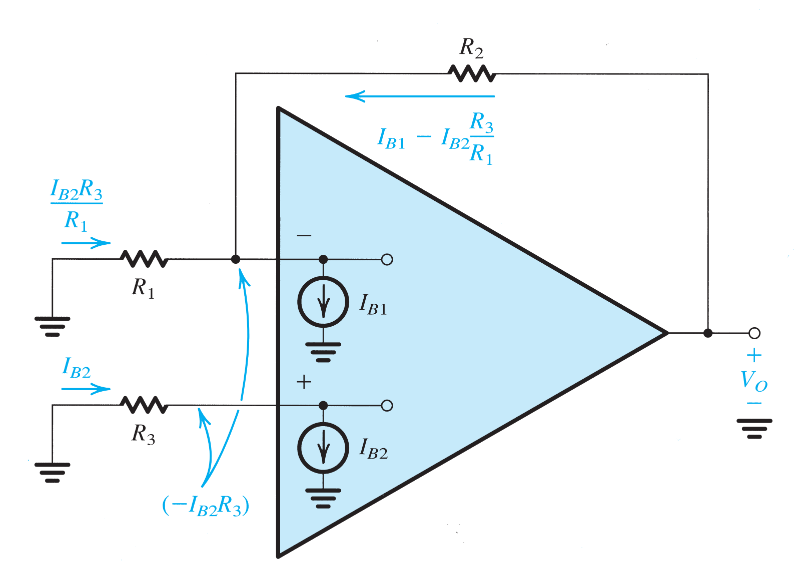


Figure 2.34 Reducing the effect of the input bias currents by introducing a resistor R_3.

This obviously places an upper limit on the value of R_2. Fortunately, however, a technique exists for reducing the value of the output dc voltage due to the input bias currents. The method consists of introducing a resistance R_3 in series with the noninverting input lead, as shown in Fig. 2.34. From a signal point of view, R_3 has a negligible effect (ideally no effect). The appropriate value for R_3 can be determined by analyzing the circuit in Fig. 2.34, where analysis details are shown, and the output voltage is given by

$$V_O = -I_{B2}R_3 + R_2(I_{B1} - I_{B2}R_3/R_1) \tag{2.38}$$

Consider first the case $I_{B1} = I_{B2} = I_B$, which results in

$$V_O = I_B[R_2 - R_3(1 + R_2/R_1)]$$

Thus we can reduce V_O to zero by selecting R_3 such that

$$R_3 = \frac{R_2}{1 + R_2/R_1} = \frac{R_1R_2}{R_1 + R_2} \tag{2.39}$$

That is, R_3 should be made equal to the parallel equivalent of R_1 and R_2.

Having selected R_3 as above, let us evaluate the effect of a finite offset current I_{OS}. Let $I_{B1} = I_B + I_{OS}/2$ and $I_{B2} = I_B - I_{OS}/2$, and substitute in Eq. (2.38). The result is

$$V_O = I_{OS}R_2 \tag{2.40}$$

which is usually about an order of magnitude smaller than the value obtained without R_3 (Eq. 2.37). We conclude that to minimize the effect of the input bias currents, one should *place in the positive lead a resistance equal to the equivalant* dc *resistance seen by the inverting terminal.* We emphasize the word *dc* in the last statement; note that if the amplifier is ac-coupled, we should select $R_3 = R_2$, as shown in Fig. 2.35.

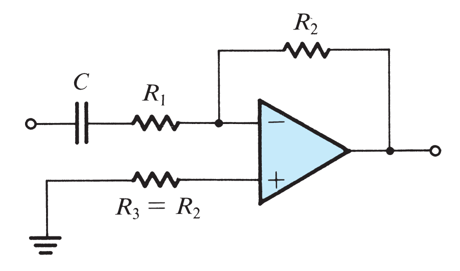


Figure 2.35 In an ac-coupled amplifier the dc resistance seen by the inverting terminal is R_2; hence R_3 is chosen equal to R_2.

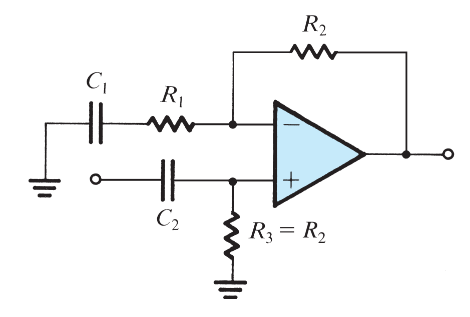


Figure 2.36 Illustrating the need for a continuous dc path for each of the op-amp input terminals. Specifically, note that the amplifier will *not* work without resistor R_3.

While we are on the subject of ac-coupled amplifiers, we should note that one must always provide a continuous dc path between each of the input terminals of the op amp and ground. This is the case no matter how small I_B is. For this reason the ac-coupled noninverting amplifier of Fig. 2.36 will *not* work without the resistance R_3 to ground. Unfortunately, including R_3 lowers considerably the input resistance of the closed-loop amplifier.

EXERCISE

2.24 Consider an inverting amplifier circuit designed using an op amp and two resistors, $R_1 = 10$ kΩ and $R_2 = 1$ MΩ. If the op amp is specified to have an input bias current of 100 nA and an input offset current of 10 nA, find the output dc offset voltage resulting and the value of a resistor R_3 to be placed in series with the positive input lead in order to minimize the output offset voltage. What is the new value of V_O?
Ans. 0.1 V; 9.9 kΩ ($\simeq$ 10 kΩ); 0.01 V

2.6.3 Effect of V_{OS} and I_{OS} on the Operation of the Inverting Integrator

Our discussion of the inverting integrator circuit in Section 2.5.2 mentioned the susceptibility of this circuit to saturation in the presence of small dc voltages or currents. It behooves us therefore to consider the effect of the op-amp dc offsets on its operation. As will be seen, these effects can be quite dramatic.

$$v_O = V_{OS} + \frac{1}{C}\int_0^t \frac{V_{OS}}{R}dt$$
$$= V_{OS} + \frac{V_{OS}}{CR}t$$

Figure 2.37 Determining the effect of the op-amp input offset voltage V_{OS} on the Miller integrator circuit. Note that since the output rises with time, the op amp eventually saturates.

To see the effect of the input dc offset voltage V_{OS}, consider the integrator circuit in Fig. 2.37, where for simplicity we have short-circuited the input signal source. Analysis of the circuit is straightforward and is shown in Fig. 2.37. Assuming for simplicity that at time $t = 0$ the voltage across the capacitor is zero, the output voltage as a function of time is given by

$$v_O = V_{OS} + \frac{V_{OS}}{CR}t \tag{2.41}$$

Thus v_O increases linearly with time until the op amp saturates—clearly an unacceptable situation! As should be expected, the dc input offset current I_{OS} produces a similar problem. Figure 2.38 illustrates the situation. Observe that we have added a resistance R in the op-amp positive-input lead in order to keep the input bias current I_B from flowing through C. Nevertheless, the offset current I_{OS} will flow through C and cause v_O to ramp linearly with time until the op amp saturates.

As mentioned in Section 2.5.2 the dc problem of the integrator circuit can be alleviated by connecting a resistor R_F across the integrator capacitor C, as shown in Fig. 2.25. Such a resistor provides a dc path through which the dc currents (V_{OS}/R) and I_{OS} can flow (assuming a resistance equal to $R \| R_F$ is connected in the positive op-amp lead), with the result that v_O will now have a dc component $[V_{OS}(1+R_F/R)+I_{OS}R_F]$ instead of rising linearly. To keep the dc offset at the output small, one would select a low value for R_F. Unfortunately, however, the lower the value of R_F, the less ideal the integrator circuit becomes.

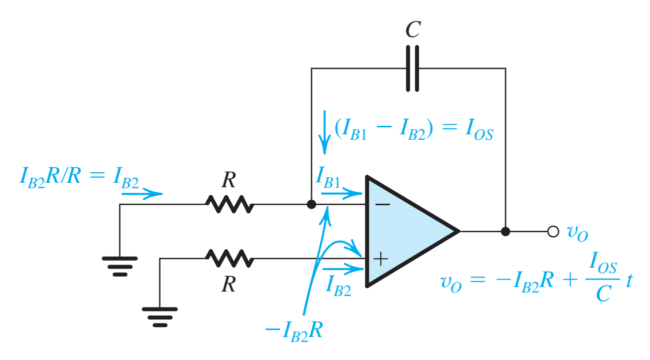


Figure 2.38 Effect of the op-amp input bias and offset currents on the performance of the Miller integrator circuit.

EXERCISE

2.25 Consider a Miller integrator with a time constant of 1 ms and an input resistance of 10 kΩ. Let the op amp have $V_{OS} = 2$ mV and output saturation voltages of ±12 V. (a) Assuming that when the power supply is turned on the capacitor voltage is zero, how long does it take for the amplifier to saturate? (b) Select the largest possible value for a feedback resistor R_F so that at least ±10 V of output signal swing remains available. What is the corner frequency of the resulting STC network?
Ans. (a) 6 s; (b) 10 MΩ, 0.16 Hz

2.7 Effect of Finite Open-Loop Gain and Bandwidth on Circuit Performance

2.7.1 Frequency Dependence of the Open-Loop Gain

The differential open-loop gain A of an op amp is not infinite; rather, it is finite and decreases with frequency. Figure 2.39 shows a plot for $|A|$, with the numbers typical of some commercially available general-purpose op amps (such as the popular 741-type op amp, available from many semiconductor manufacturers; its internal circuit is studied in Chapter 13).

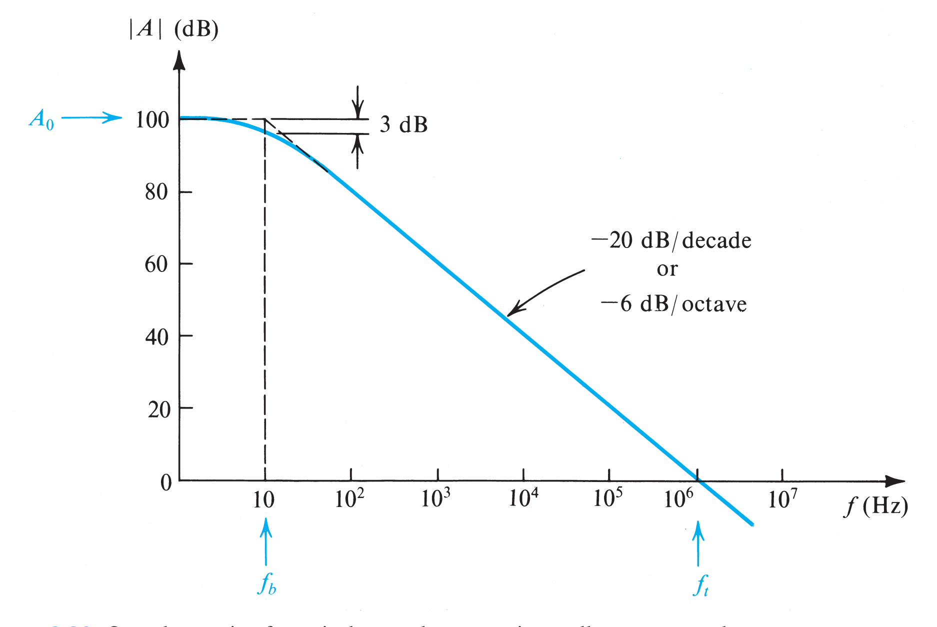


Figure 2.39 Open-loop gain of a typical general-purpose, internally compensated op amp.

Note that although the gain is quite high at dc and low frequencies, it starts to fall off at a rather low frequency (10 Hz in our example). The uniform –20-dB/decade gain rolloff shown is typical of **internally compensated** op amps. These are units that have a network (usually a single capacitor) included within the same IC chip whose function is to cause the op-amp gain to have the single-time-constant (STC) low-pass response shown. This process of modifying the open-loop gain is termed **frequency compensation**, and its purpose is to ensure that op-amp circuits will be stable (as opposed to oscillatory). The subject of stability of op-amp circuits—or, more generally, of feedback amplifiers—will be studied in Chapter 11.

By analogy to the response of low-pass STC circuits (see Section 1.6 and, for more detail, Appendix E), the gain $A(s)$ of an internally compensated op amp may be expressed as

$$A(s) = \frac{A_0}{1 + s/\omega_b} \tag{2.42}$$

which for physical frequencies, $s = j\omega$, becomes

$$A(j\omega) = \frac{A_0}{1 + j\omega/\omega_b} \tag{2.43}$$

where A_0 denotes the dc gain and ω_b is the 3-dB frequency (corner frequency or "break" frequency). For the example shown in Fig. 2.39, $A_0 = 10^5$ and $\omega_b = 2\pi \times 10$ rad/s. For frequencies $\omega \gg \omega_b$ (about 10 times and higher) Eq. (2.43) may be approximated by

$$A(j\omega) \simeq \frac{A_0\omega_b}{j\omega} \tag{2.44}$$

Thus,

$$|A(j\omega)| = \frac{A_0\omega_b}{\omega} \tag{2.45}$$

from which it can be seen that the gain $|A|$ reaches unity (0 dB) at a frequency denoted by ω_t and given by

$$\omega_t = A_0\omega_b \tag{2.46}$$

Substituting in Eq. (2.44) gives

$$A(j\omega) \simeq \frac{\omega_t}{j\omega} \tag{2.47}$$

The frequency $f_t = \omega_t/2\pi$ is usually specified on the data sheets of commercially available op amps and is known as the **unity-gain bandwidth**.[6] Also note that for $\omega \gg \omega_b$ the open-loop gain in Eq. (2.42) becomes

$$A(s) \simeq \frac{\omega_t}{s} \tag{2.48}$$

[6]Since f_t is the product of the dc gain A_0 and the 3-dB bandwidth f_b (where $f_b = \omega_b/2\pi$), it is also known as the **gain–bandwidth product** (GB). The reader is cautioned, however, that in some amplifiers (those that do not have an STC response), the unity-gain frequency and the gain–bandwidth product are *not* equal.

The gain magnitude can be obtained from Eq. (2.47) as

$$|A(j\omega)| \simeq \frac{\omega_t}{\omega} = \frac{f_t}{f} \tag{2.49}$$

Thus if f_t is known (10^6 Hz in our example), one can easily determine the magnitude of the op-amp gain at a given frequency f. Furthermore, observe that this relationship correlates with the Bode plot in Fig. 2.39. Specifically, for $f \gg f_b$, doubling f (an octave increase) results in halving the gain (a 6-dB reduction). Similarly, increasing f by a factor of 10 (a decade increase) results in reducing $|A|$ by a factor of 10 (20 dB).

As a matter of practical importance, we note that the production spread in the value of f_t between op-amp units of the same type is usually much smaller than that observed for A_0 and f_b. For this reason f_t is preferred as a specification parameter. Finally, it should be mentioned that an op amp having this uniform –6-dB/octave (or equivalently –20-dB/decade) gain rolloff is said to have a **single-pole model**. Also, since this single pole *dominates* the amplifier frequency response, it is called a *dominant pole*. For more on poles (and zeros), the reader may wish to consult Appendix F.

EXERCISE

2.26 An internally compensated op amp is specified to have an open-loop dc gain of 106 dB and a unity-gain bandwidth of 3 MHz. Find f_b and the open-loop gain (in dB) at f_b, 300 Hz, 3 kHz, 12 kHz, and 60 kHz.
Ans. 15 Hz; 103 dB; 80 dB; 60 dB; 48 dB; 34 dB

2.7.2 Frequency Response of Closed-Loop Amplifiers

We next consider the effect of limited op-amp gain and bandwidth on the closed-loop transfer functions of the two basic configurations: the inverting circuit of Fig. 2.5 and the noninverting circuit of Fig. 2.12. The closed-loop gain of the inverting amplifier, assuming a finite op-amp open-loop gain A, was derived in Section 2.2 and given in Eq. (2.5), which we repeat here as

$$\frac{V_o}{V_i} = \frac{-R_2/R_1}{1+(1+R_2/R_1)/A} \tag{2.50}$$

Substituting for A from Eq. (2.42) and using Eq. (2.46) gives

$$\frac{V_o(s)}{V_i(s)} = \frac{-R_2/R_1}{1+\dfrac{1}{A_0}\left(1+\dfrac{R_2}{R_1}\right)+\dfrac{s}{\omega_t/(1+R_2/R_1)}} \tag{2.51}$$

For $A_0 \gg 1+R_2/R_1$, which is usually the case,

$$\frac{V_o(s)}{V_i(s)} \simeq \frac{-R_2/R_1}{1+\dfrac{s}{\omega_t/(1+R_2/R_1)}} \tag{2.52}$$

which is of the same form as that for a low-pass STC network (see Table 1.2, page 36). Thus the inverting amplifier has an STC low-pass response with a dc gain of magnitude equal to R_2/R_1. The closed-loop gain rolls off at a uniform –20-dB/decade slope with a corner frequency (3-dB frequency) given by

$$\omega_{3\text{dB}} = \frac{\omega_t}{1 + R_2/R_1} \tag{2.53}$$

Similarly, analysis of the noninverting amplifier of Fig. 2.12, assuming a finite open-loop gain A, yields the closed-loop transfer function

$$\frac{V_o}{V_i} = \frac{1 + R_2/R_1}{1 + (1 + R_2/R_1)/A} \tag{2.54}$$

Substituting for A from Eq. (2.42) and making the approximation $A_0 \gg 1 + R_2/R_1$ results in

$$\frac{V_o(s)}{V_i(s)} \simeq \frac{1 + R_2/R_1}{1 + \dfrac{s}{\omega_t/(1 + R_2/R_1)}} \tag{2.55}$$

Thus the noninverting amplifier has an STC low-pass response with a dc gain of $(1 + R_2/R_1)$ and a 3-dB frequency given also by Eq. (2.53).

Example 2.6

Consider an op amp with $f_t = 1$ MHz. Find the 3-dB frequency of closed-loop amplifiers with nominal gains of +1000, +100, +10, +1, −1, −10, −100, and −1000. Sketch the magnitude frequency response for the amplifiers with closed-loop gains of +10 and −10.

Solution

We use Eq. (2.53) to obtain the results given in the following table.

Closed-Loop Gain	R_2/R_1	$f_{3\text{ dB}} = f_t/(1 + R_2/R_1)$
+1000	999	1 kHz
+100	99	10 kHz
+10	9	100 kHz
+ 1	0	1 MHz
− 1	1	0.5 MHz
−10	10	90.9 kHz
−100	100	9.9 kHz
−1000	1000	≃ 1 kHz

Figure 2.40 shows the frequency response for the amplifier whose nominal dc gain is +10 (20 dB), and Fig. 2.41 shows the frequency response for the –10 (also 20 dB) case. An interesting observation follows

from the table above: The unity-gain inverting amplifier has a 3-dB frequency of $f_t/2$ as compared to f_t for the unity-gain noninverting amplifier (the unity-gain voltage follower).

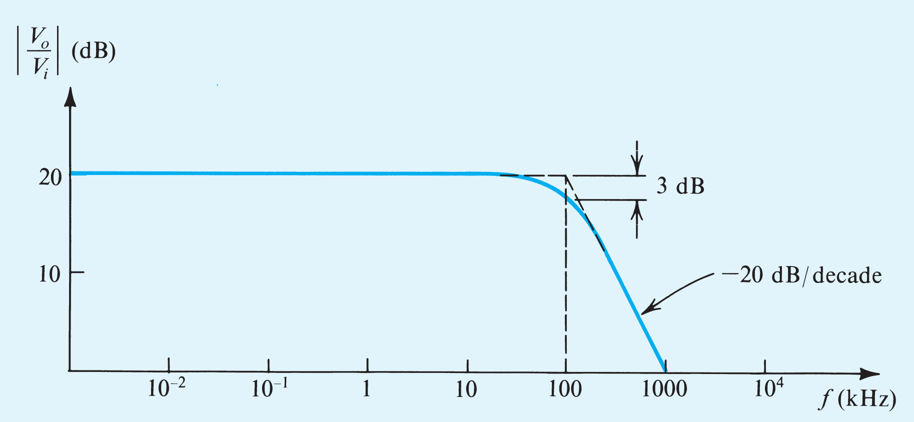


Figure 2.40 Frequency response of an amplifier with a nominal gain of +10 V/V.

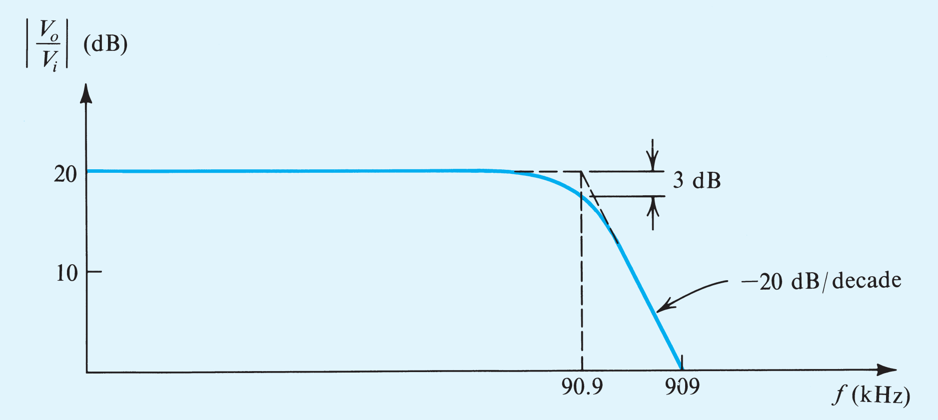


Figure 2.41 Frequency response of an amplifier with a nominal gain of −10 V/V.

The table in Example 2.6 above clearly illustrates the trade-off between gain and bandwidth: For a given op amp, the lower the closed-loop gain required, the wider the bandwidth achieved. Indeed, the noninverting configuration exhibits a constant **gain–bandwidth product** equal to f_t of the op amp. An interpretation of these results in terms of feedback theory will be given in Chapter 11.

EXERCISES

2.27 An internally compensated op amp has a dc open-loop gain of 10^6 V/V and an open-loop gain of 40 dB at 10 kHz. Estimate its 3-dB frequency, its unity-gain frequency, its gain–bandwidth product, and its expected gain at 1 kHz.
Ans. 1 Hz; 1 MHz; 1 MHz; 60 dB

2.28 An op amp having a 106-dB gain at dc and a single-pole frequency response with $f_t = 2$ MHz is used to design a noninverting amplifier with nominal dc gain of 100. Find the 3-dB frequency of the closed-loop gain.
Ans. 20 kHz

2.8 Large-Signal Operation of Op Amps

In this section, we study the limitations on the performance of op-amp circuits when large output signals are present.

2.8.1 Output Voltage Saturation

Similar to all other amplifiers, op amps operate linearly over a limited range of output voltages. Specifically, the op-amp output saturates in the manner shown in Fig. 1.14 with L_+ and L_- within 1 V or so of the positive and negative power supplies, respectively. Thus, an op amp that is operating from ±15-V supplies will saturate when the output voltage reaches about +13 V in the positive direction and –13 V in the negative direction. For this particular op amp the **rated output voltage** is said to be ±13 V. To avoid clipping off the peaks of the output waveform, and the resulting waveform distortion, the input signal must be kept correspondingly small.

2.8.2 Output Current Limits

Another limitation on the operation of op amps is that their output current is limited to a specified maximum. For instance, the popular 741 op amp is specified to have a maximum output current of ±20 mA. Thus, in designing closed-loop circuits utilizing the 741, the designer has to ensure that under no condition will the op amp be required to supply an output current, in either direction, exceeding 20 mA. This, of course, has to include both the current in the feedback circuit as well as the current supplied to a load resistor. If the circuit requires a larger current, the op-amp output voltage will saturate at the level corresponding to the maximum allowed output current.

Example 2.7

Consider the noninverting amplifier circuit shown in Fig. 2.42. As shown, the circuit is designed for a nominal gain $(1 + R_2/R_1) = 10$ V/V. It is fed with a low-frequency sine-wave signal of peak voltage V_p and is connected to a load resistor R_L. The op amp is specified to have output saturation voltages of ±13 V and output current limits of ±20 mA.

(a) For $V_p = 1$ V and $R_L = 1$ kΩ, specify the signal resulting at the output of the amplifier.
(b) For $V_p = 1.5$ V and $R_L = 1$ kΩ, specify the signal resulting at the output of the amplifier.
(c) For $R_L = 1$ kΩ, what is the maximum value of V_p for which an undistorted sine-wave output is obtained?
(d) For $V_p = 1$ V, what is the lowest value of R_L for which an undistorted sine-wave output is obtained?

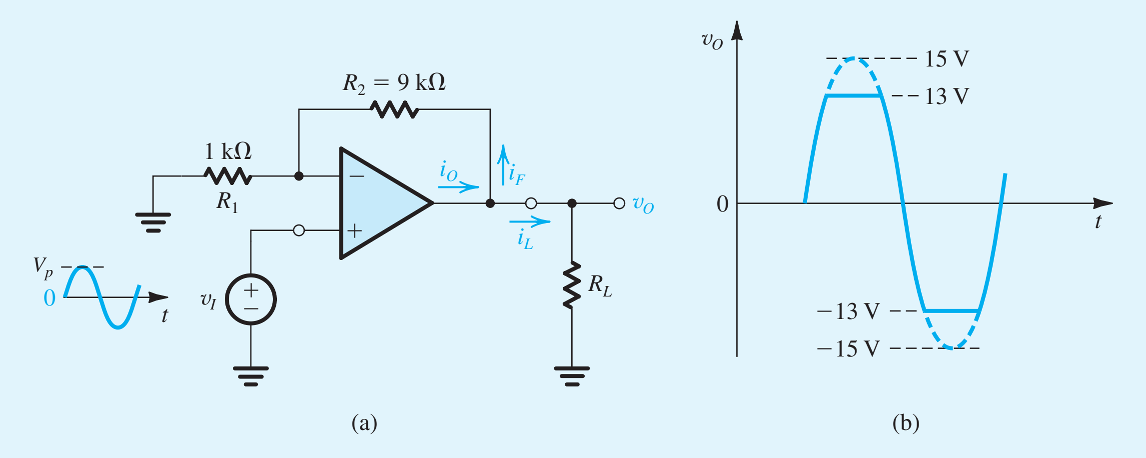


Figure 2.42 **(a)** A noninverting amplifier with a nominal gain of 10 V/V designed using an op amp that saturates at ±13-V output voltage and has ±20-mA output current limits. **(b)** When the input sine wave has a peak of 1.5 V, the output is clipped off at ±13 V.

Solution

(a) For $V_p = 1$ V and $R_L = 1$ kΩ, the output will be a sine wave with peak value of 10 V. This is lower than output saturation levels of ±13 V, and thus the amplifier is not limited that way. Also, when the output is at its peak (10 V), the current in the load will be 10 V/1 kΩ = 10 mA, and the current in the feedback network will be 10 V/(9 + 1) kΩ = 1 mA, for a total op-amp output current of 11 mA, well under its limit of 20 mA.

(b) Now if V_p is increased to 1.5 V, ideally the output would be a sine wave of 15-V peak. The op amp, however, will saturate at ±13 V, thus clipping the sine-wave output at these levels. Let's next check on the op-amp output current: At 13-V output and $R_L = 1$ kΩ, $i_L = 13$ mA and $i_F = 1.3$ mA; thus $i_O = 14.3$ mA, again under the 20-mA limit. Thus the output will be a sine wave with its peaks clipped off at ±13 V, as shown in Fig. 2.42(b).

(c) For $R_L = 1$ kΩ, the maximum value of V_p for undistorted sine-wave output is 1.3 V. The output will be a 13-V peak sine wave, and the op-amp output current at the peaks will be 14.3 mA.

Example 2.7 *continued*

(d) For $V_p = 1$ V and R_L reduced, the lowest value possible for R_L while the output is remaining an undistorted sine wave of 10-V peak can be found from

$$i_{O\max} = 20 \text{ mA} = \frac{10 \text{ V}}{R_{L\min}} + \frac{10 \text{ V}}{9 \text{ k}\Omega + 1 \text{ k}\Omega}$$

which results in

$$R_{L\min} = 526 \ \Omega$$

2.8.3 Slew Rate

Another phenomenon that can cause nonlinear distortion when large output signals are present is slew-rate limiting. The name refers to the fact that there is a specific *maximum rate of change* possible at the output of a real op amp. This maximum is known as the **slew rate** (SR) of the op amp and is defined as

$$\text{SR} = \left.\frac{dv_O}{dt}\right|_{\max} \tag{2.56}$$

and is usually specified on the op-amp data sheet in units of V/μs. It follows that if the input signal applied to an op-amp circuit is such that it demands an output response that is faster than the specified value of SR, the op amp will not comply. Rather, its output will change at the maximum possible rate, which is equal to its SR. As an example, consider an op amp connected in the unity-gain voltage-follower configuration shown in Fig. 2.43(a), and let the input signal be the step voltage shown in Fig. 2.43(b). The output of the op amp will not be able to rise instantaneously to the ideal value V; rather, the output will be the linear ramp of slope equal to SR, shown in Fig. 2.43(c). The amplifier is then said to be **slewing**, and its output is **slew-rate limited**.

In order to understand the origin of the slew-rate phenomenon, we need to know about the internal circuit of the op amp, and we will study it in Chapter 13. For the time being, however, it is sufficient to know about the phenomenon and to note that it is distinct from the finite op-amp bandwidth that limits the frequency response of the closed-loop amplifiers, studied in the previous section. The limited bandwidth is a linear phenomenon and does not result in a change in the shape of an input sinusoid; that is, it does not lead to nonlinear distortion. The slew-rate limitation, on the other hand, can cause nonlinear distortion to an input sinusoidal signal when its frequency and amplitude are such that the corresponding ideal output would require v_O to change at a rate greater than SR. This is the origin of another related op-amp specification, its full-power bandwidth, to be explained later.

Before leaving the example in Fig. 2.43, however, we should point out that if the step input voltage V is sufficiently small, the output can be the exponentially rising ramp shown

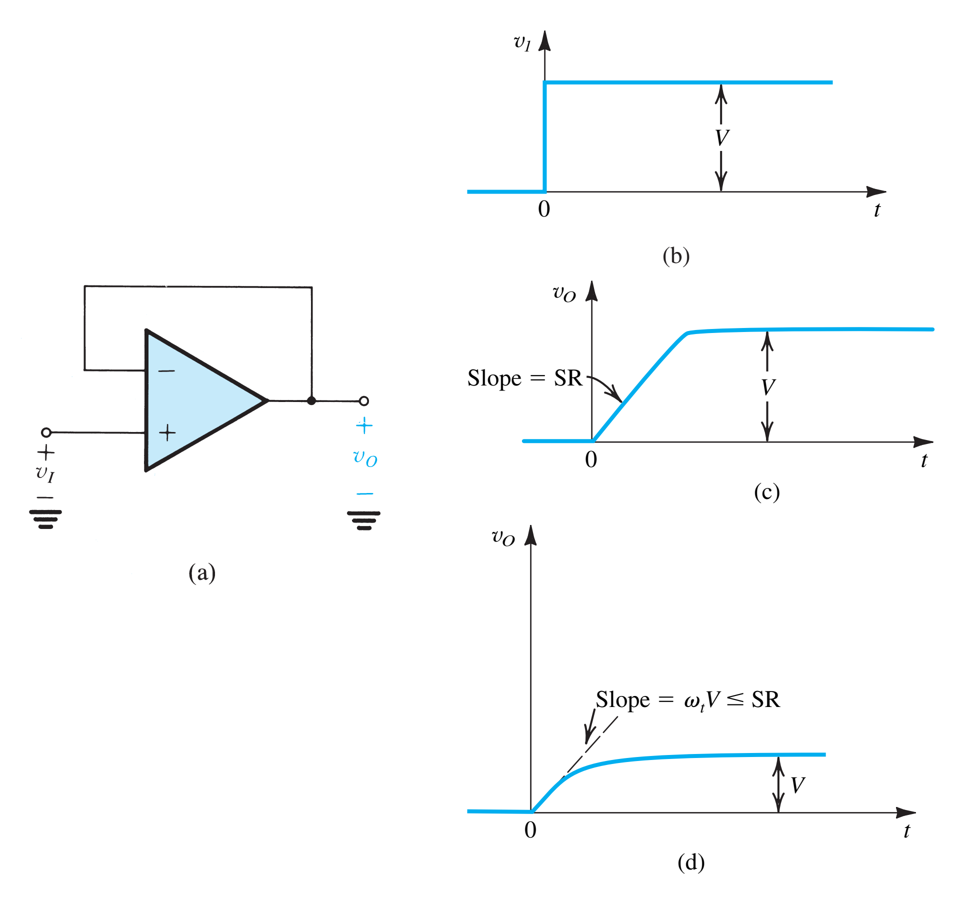


Figure 2.43 **(a)** Unity-gain follower. **(b)** Input step waveform. **(c)** Linearly rising output waveform obtained when the amplifier is slew-rate limited. **(d)** Exponentially rising output waveform obtained when V is sufficiently small so that the initial slope $(\omega_t V)$ is smaller than or equal to SR.

in Fig. 2.43(d). Such an output would be expected from the follower if the only limitation on its dynamic performance were the finite op-amp bandwidth. Specifically, the transfer function of the follower can be found by substituting $R_1 = \infty$ and $R_2 = 0$ in Eq. (2.55) to obtain

$$\frac{V_o}{V_i} = \frac{1}{1 + s/\omega_t} \tag{2.57}$$

which is a low-pass STC response with a time constant $1/\omega_t$. Its step response would therefore be (see Appendix E)

$$v_O(t) = V(1 - e^{-\omega_t t}) \tag{2.58}$$

The initial slope of this exponentially rising function is $(\omega_t V)$. Thus, as long as V is sufficiently small so that $\omega_t V \leq \text{SR}$, the output will be as in Fig. 2.43(d).

EXERCISE

2.29 An op amp that has a slew rate of 1 V/μs and a unity-gain bandwidth f_t of 1 MHz is connected in the unity-gain follower configuration. Find the largest possible input voltage step for which the output waveform will still be given by the exponential ramp of Eq. (2.58). For this input voltage, what is the 10% to 90% rise time of the output waveform? If an input step 10 times as large is applied, find the 10% to 90% rise time of the output waveform.
Ans. 0.16 V; 0.35 μs; 1.28 μs

2.8.4 Full-Power Bandwidth

Op-amp slew-rate limiting can cause nonlinear distortion in sinusoidal waveforms. Consider once more the unity-gain follower with a sine-wave input given by

$$v_I = \hat{V}_i \sin \omega t$$

The rate of change of this waveform is given by

$$\frac{dv_I}{dt} = \omega \hat{V}_i \cos \omega t$$

with a maximum value of $\omega \hat{V}_i$. This maximum occurs at the zero crossings of the input sinusoid. Now if $\omega \hat{V}_i$ exceeds the slew rate of the op amp, the output waveform will be distorted in the manner shown in Fig. 2.44. Observe that the output cannot keep up with the large rate of change of the sinusoid at its zero crossings, and the op amp slews.

The op-amp data sheets usually specify a frequency f_M called the **full-power bandwidth**. It is the frequency at which an output sinusoid with amplitude equal to the rated output voltage of the op amp begins to show distortion due to slew-rate limiting. If we denote the rated output

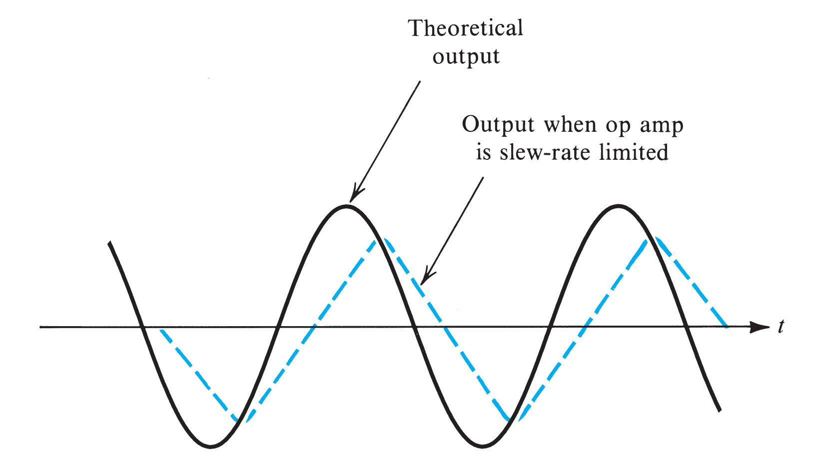


Figure 2.44 Effect of slew-rate limiting on output sinusoidal waveforms.

voltage $V_{o\max}$, then f_M is related to SR as follows:

$$\omega_M V_{o\max} = \text{SR}$$

Thus,

$$f_M = \frac{\text{SR}}{2\pi V_{o\max}} \tag{2.59}$$

It should be obvious that output sinusoids of amplitudes smaller than $V_{o\max}$ will show slew-rate distortion at frequencies higher than ω_M. In fact, at a frequency ω higher than ω_M, the maximum amplitude of the undistorted output sinusoid is given by

$$V_o = V_{o\max}\left(\frac{\omega_M}{\omega}\right) \tag{2.60}$$

EXERCISE

2.30 An op amp has a rated output voltage of ±10 V and a slew rate of 1 V/μs. What is its full-power bandwidth? If an input sinusoid with frequency $f = 5f_M$ is applied to a unity-gain follower constructed using this op amp, what is the maximum possible amplitude that can be accommodated at the output without incurring SR distortion?
Ans. 15.9 kHz; 2 V (peak)

Summary

- The IC op amp is a versatile circuit building block. It is easy to apply, and the performance of op-amp circuits closely matches theoretical predictions.
- The op-amp terminals are the inverting input terminal (1), the noninverting input terminal (2), the output terminal (3), the positive-supply terminal (4) to be connected to the positive power supply (V_{CC}), and the negative-supply terminal (5) to be connected to the negative supply ($-V_{EE}$). The common terminal of the two supplies is the circuit ground.
- The ideal op amp responds only to the difference input signal, that is, $(v_2 - v_1)$; it provides at the output, between terminal 3 and ground, a signal $A(v_2 - v_1)$, where A, the open-loop gain, is very large (10^4 to 10^6) and ideally infinite; and it has an infinite input resistance and a zero output resistance. (See Table 2.1.)
- Negative feedback is applied to an op amp by connecting a passive component between its output terminal and its inverting (negative) input terminal. Negative feedback causes the voltage between the two input terminals to become very small and ideally zero. Correspondingly, a virtual short circuit is said to exist between the two input terminals. If the positive input terminal is connected to ground, a virtual ground appears on the negative input terminal.
- The two most important assumptions in the analysis of op-amp circuits, presuming negative feedback exists and the op amps are ideal, are as follows: the two input terminals of the op amp are at the same voltage, and zero current flows into the op-amp input terminals.
- With negative feedback applied and the loop closed, the closed-loop gain is almost entirely determined by external components: For the inverting configuration, $V_o/V_i = -R_2/R_1$; and for the noninverting configuration, $V_o/V_i = 1 + R_2/R_1$.
- The noninverting closed-loop configuration features a very high input resistance. A special case is the unity-gain

follower, frequently employed as a buffer amplifier to connect a high-resistance source to a low-resistance load.

- The difference amplifier of Fig. 2.16 is designed with $R_4/R_3 = R_2/R_1$, resulting in $v_O = (R_2/R_1)(v_{I2} - v_{I1})$.
- The instrumentation amplifier of Fig. 2.20(b) is a very popular circuit. It provides $v_O = (1+R_2/R_1)(R_4/R_3)(v_{I2} - v_{I1})$. It is usually designed with $R_3 = R_4$, and R_1 and R_2 selected to provide the required gain. If an adjustable gain is needed, part of R_1 can be made variable.
- The inverting Miller integrator of Fig. 2.24(a) is a popular circuit, frequently employed in analog signal-processing functions such as filters (Chapter 17) and oscillators (Chapter 18).
- The input offset voltage, V_{OS}, is the magnitude of dc voltage that when applied between the op-amp input terminals, with appropriate polarity, reduces the dc offset voltage at the output to zero.
- The effect of V_{OS} on performance can be evaluated by including in the analysis a dc source V_{OS} in series with the op-amp positive input lead. For both the inverting and the noninverting configurations, V_{OS} results in a dc offset voltage at the output of $V_{OS}(1 + R_2/R_1)$.
- Capacitively coupling an op amp reduces the dc offset voltage at the output considerably.
- The average of the two dc currents, I_{B1} and I_{B2}, that flow in the input terminals of the op amp, is called the input bias current, I_B. In a closed-loop amplifier, I_B gives rise to a dc offset voltage at the output of magnitude $I_B R_2$. This voltage can be reduced to $I_{OS} R_2$ by connecting a resistance in series with the positive input terminal equal to the total dc resistance seen by the negative input terminal. I_{OS} is the input offset current; that is, $I_{OS} = |I_{B1} - I_{B2}|$.
- Connecting a large resistance in parallel with the capacitor of an op-amp inverting integrator prevents op-amp saturation (due to the effect of V_{OS} and I_B).
- For most internally compensated op amps, the open-loop gain falls off with frequency at a rate of −20 dB/decade, reaching unity at a frequency f_t (the unity-gain bandwidth). Frequency f_t is also known as the gain–bandwidth product of the op amp: $f_t = A_0 f_b$, where A_0 is the dc gain, and f_b is the 3-dB frequency of the open-loop gain. At any frequency f ($f \gg f_b$), the op-amp gain $|A| \simeq f_t/f$.
- For both the inverting and the noninverting closed-loop configurations, the 3-dB frequency is equal to $f_t/(1 + R_2/R_1)$.
- The maximum rate at which the op-amp output voltage can change is called the slew rate. The slew rate, SR, is usually specified in V/μs. Op-amp slewing can result in nonlinear distortion of output signal waveforms.
- The full-power bandwidth, f_M, is the maximum frequency at which an output sinusoid with an amplitude equal to the op-amp rated output voltage ($V_{o\max}$) can be produced without distortion: $f_M = \text{SR}/2\pi V_{o\max}$.

PROBLEMS

Computer Simulation Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as allowable signal swing and amplifier nonlinear distortion. Instructions to assist in setting up PSPice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 2.1: The Ideal Op Amp

2.1 What is the minimum number of pins required for a so-called dual-op-amp IC package, one containing two op amps? What is the number of pins required for a so-called quad-op-amp package, one containing four op-amps?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

2.2 The circuit of Fig. P2.2 uses an op amp that is ideal except for having a finite gain A. Measurements indicate $v_O = 4.0$ V when $v_I = 1.0$ V. What is the op-amp gain A?

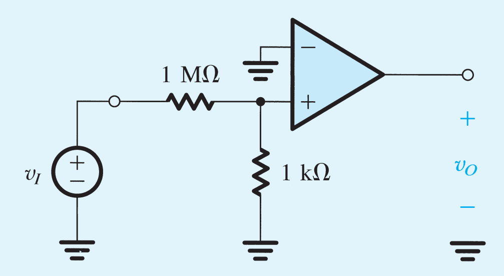


Figure P2.2

2.3 Measurement of a circuit incorporating what is thought to be an ideal op amp shows the voltage at the op-amp output to be −2.000 V and that at the negative input to be −1.000 V. For the amplifier to be ideal, what would you expect the voltage at the positive input to be? If the measured voltage at the positive input is −1.005 V, what is likely to be the actual gain of the amplifier?

2.4 A set of experiments is run on an op amp that is ideal except for having a finite gain A. The results are tabulated below. Are the results consistent? If not, are they reasonable, in view of the possibility of experimental error? What do they show the gain to be? Using this value, predict values of the measurements that were accidentally omitted (the blank entries).

Experiment #	v_1	v_2	v_O
1	0.00	0.00	0.00
2	1.00	1.00	0.00
3		1.00	1.00
4	1.00	1.10	10.1
5	2.01	2.00	−0.99
6	1.99	2.00	1.00
7	5.10		−5.10

2.5 Refer to Exercise 2.3. This problem explores an alternative internal structure for the op amp. In particular, we wish to model the internal structure of a particular op amp using two transconductance amplifiers and one transresistance amplifier. Suggest an appropriate topology. For equal transconductances G_m and a transresistance R_m, find an expression for the open-loop gain A. For $G_m = 40$ mA/V and $R_m = 1 \times 10^6\ \Omega$, what value of A results?

2.6 The two wires leading from the output terminals of a transducer pick up an interference signal that is a 60-Hz, 2-V sinusoid. The output signal of the transducer is sinusoidal of 5-mV amplitude and 1000-Hz frequency. Give expressions for v_{cm}, v_d, and the total signal between each wire and the system ground.

2.7 Nonideal (i.e., real) operational amplifiers respond to both the differential and common-mode components of their input signals (refer to Fig. 2.4 for signal representation). Thus the output voltage of the op amp can be expressed as

$$v_O = A_d v_{Id} + A_{cm} v_{Icm}$$

where A_d is the differential gain (referred to simply as A in the text) and A_{cm} is the common-mode gain (assumed to be zero in the text). The op amp's effectiveness in rejecting common-mode signals is measured by its CMRR, defined as

$$\text{CMRR} = 20 \log \left| \frac{A_d}{A_{cm}} \right|$$

Consider an op amp whose internal structure is of the type shown in Fig. E2.3 except for a mismatch ΔG_m between the transconductances of the two channels; that is,

$$G_{m1} = G_m - \tfrac{1}{2}\Delta G_m$$

$$G_{m2} = G_m + \tfrac{1}{2}\Delta G_m$$

Find expressions for A_d, A_{cm}, and CMRR. What is the maximum permitted percentage mismatch between the two G_m values if a minimum CMRR of 60 dB is required?

Section 2.2: The Inverting Configuration

2.8 Assuming ideal op amps, find the voltage gain v_o/v_i and input resistance R_{in} of each of the circuits in Fig. P2.8.

2.9 A particular inverting circuit uses an ideal op amp and two 10-kΩ resistors. What closed-loop gain would you expect? If a dc voltage of +1.00 V is applied at the input, what outputs result? If the 10-kΩ resistors are said to be "1% resistors," having values somewhere in the range (1 ± 0.01) times the nominal value, what range of outputs would you expect to actually measure for an input of precisely 1.00 V?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

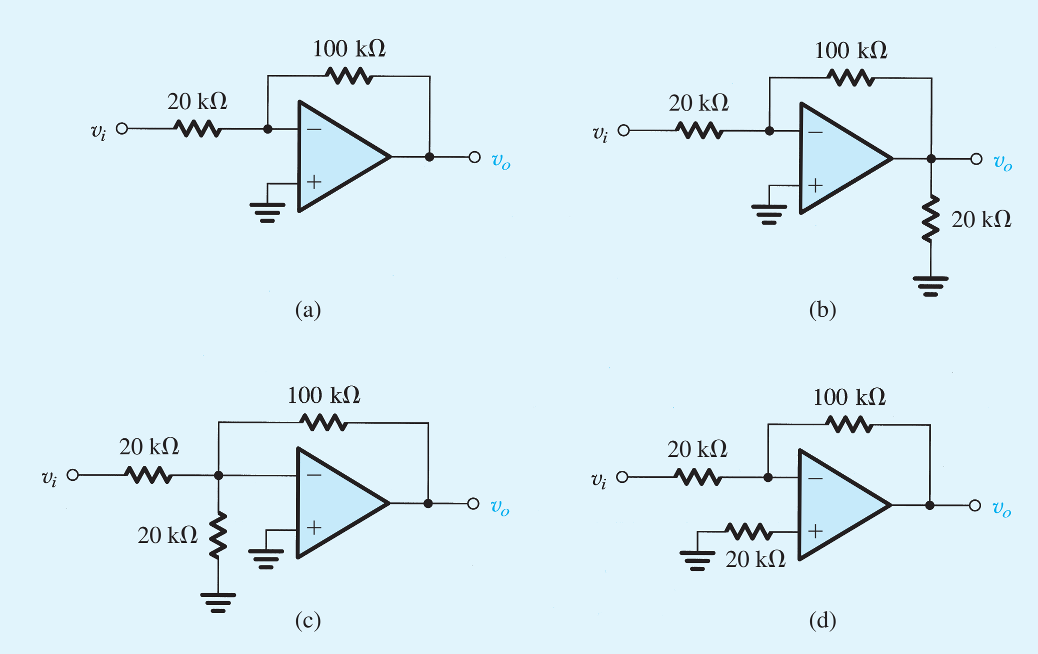


Figure P2.8

2.10 You are provided with an ideal op amp and three 10-kΩ resistors. Using series and parallel resistor combinations, how many different inverting-amplifier circuit topologies are possible? What is the largest (noninfinite) available voltage gain magnitude? What is the smallest (nonzero) available gain magnitude? What are the input resistances in these two cases?

SIM **2.11** For ideal op amps operating with the following feedback networks in the inverting configuration, what closed-loop gain results?

(a) $R_1 = 10\text{ k}\Omega, R_2 = 10\text{ k}\Omega$
(b) $R_1 = 10\text{ k}\Omega, R_2 = 100\text{ k}\Omega$
(c) $R_1 = 10\text{ k}\Omega, R_2 = 1\text{ k}\Omega$
(d) $R_1 = 100\text{ k}\Omega, R_2 = 10\text{ M}\Omega$
(e) $R_1 = 100\text{ k}\Omega, R_2 = 1\text{ M}\Omega$

D 2.12 Given an ideal op amp, what are the values of the resistors R_1 and R_2 to be used to design amplifiers with the closed-loop gains listed below? In your designs, use at least one 10-kΩ resistor and another equal or larger resistor.

(a) −1 V/V
(b) −2 V/V
(c) −5 V/V
(d) −100 V/V

D 2.13 Design an inverting op-amp circuit for which the gain is −10 V/V and the total resistance used is 110 kΩ.

D 2.14 Using the circuit of Fig. 2.5 and assuming an ideal op amp, design an inverting amplifier with a gain of 46 dB having the largest possible input resistance under the constraint of having to use resistors no larger than 1 MΩ. What is the input resistance of your design?

2.15 An ideal op amp is connected as shown in Fig. 2.5 with $R_1 = 10\text{ k}\Omega$ and $R_2 = 100\text{ k}\Omega$. A symmetrical square-wave signal with levels of 0 V and −1 V is applied at the input. Sketch and clearly label the waveform of the resulting output voltage. What is its average value? What is its highest value? What is its lowest value?

2.16 For the circuit in Fig. P2.16, assuming an ideal op amp, find the currents through all branches and the voltages at all nodes. Since the current supplied by the op amp is greater than the current drawn from the input signal source, where does the additional current come from?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

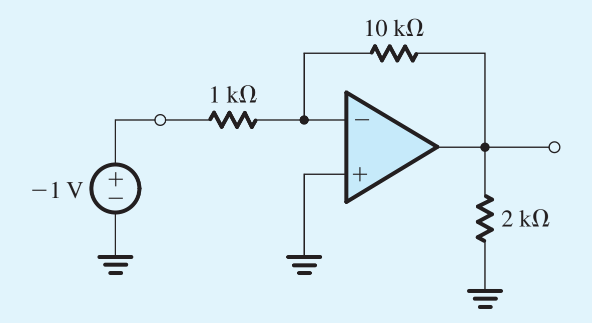


Figure P2.16

2.17 An inverting op-amp circuit is fabricated with the resistors R_1 and R_2 having $x\%$ tolerance (i.e., the value of each resistance can deviate from the nominal value by as much as $\pm x\%$). What is the tolerance on the realized closed-loop gain? Assume the op amp to be ideal. If the nominal closed-loop gain is -100 V/V and $x = 1$, what is the range of gain values expected from such a circuit?

2.18 An ideal op amp with 5-kΩ and 15-kΩ resistors is used to create a +5-V supply from a −15-V reference. Sketch the circuit. What are the voltages at the ends of the 5-kΩ resistor? If these resistors are so-called 1% resistors, whose actual values are the range bounded by the nominal value ±1%, what are the limits of the output voltage produced? If the −15-V supply can also vary by ±1%, what is the range of the output voltages that might be found?

2.19 An inverting op-amp circuit for which the required gain is −50 V/V uses an op amp whose open-loop gain is only 500 V/V. If the larger resistor used is 100 kΩ, to what must the smaller be adjusted? With what resistor must a 2-kΩ resistor connected to the input be shunted to achieve this goal? (Note that a resistor R_a is said to be shunted by resistor R_b when R_b is placed in parallel with R_a.)

D 2.20 (a) Design an inverting amplifier with a closed-loop gain of −200 V/V and an input resistance of 1 kΩ.
(b) If the op amp is known to have an open-loop gain of 5000 V/V, what do you expect the closed-loop gain of your circuit to be (assuming the resistors have precise values)?
(c) Give the value of a resistor you can place in parallel (shunt) with R_1 to restore the closed-loop gain to its nominal value. Use the closest standard 1% resistor value (see Appendix J).

2.21 An op amp with an open-loop gain of 5000 V/V is used in the inverting configuration. If in this application the output voltage ranges from −10 V to +10 V, what is the maximum voltage by which the "virtual ground node" departs from its ideal value?

2.22 The circuit in Fig. P2.22 is frequently used to provide an output voltage v_o proportional to an input signal current i_i.

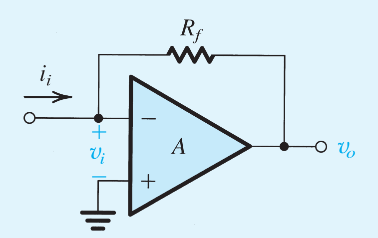


Figure P2.22

Derive expressions for the transresistance $R_m \equiv v_o/i_i$ and the input resistance $R_i \equiv v_i/i_i$ for the following cases:

(a) A is infinite.
(b) A is finite.

2.23 Show that for the inverting amplifier if the op-amp gain is A, the input resistance is given by

$$R_{in} = R_1 + \frac{R_2}{A+1}$$

2.24 For an inverting amplifier with nominal closed-loop gain R_2/R_1, find the minimum value that the op-amp open-loop gain A must have (in terms of R_2/R_1) so that the gain error (due to the finite A) is limited to 0.1%, 1%, and 10%. In each case find the value of a resistor R_{Ia} such that when it is placed in shunt with R_1, the gain is restored to its nominal value.

***2.25** Figure P2.25 shows an op amp that is ideal except for having a finite open-loop gain and is used to realize an inverting amplifier whose gain has a nominal magnitude $G = R_2/R_1$. To compensate for the gain reduction due to

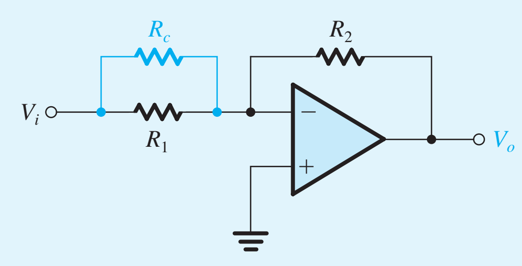


Figure P2.25

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

the finite A, a resistor R_c is shunted across R_1. Show that perfect compensation is achieved when R_c is selected according to

$$\frac{R_c}{R_1} = \frac{A - G}{1 + G}$$

D *2.26 (a) Use Eq. (2.5) to obtain the amplifier open-loop gain A required to realize a specified closed-loop gain ($G_{\text{nominal}} = -R_2/R_1$) within a specified gain error ϵ,

$$\epsilon \equiv \left| \frac{G - G_{\text{nominal}}}{G_{\text{nominal}}} \right|$$

(b) Design an inverting amplifer for a nominal closed-loop gain of -100, an input resistance of 1 kΩ, and a gain error of $\leq 10\%$. Specify R_1, R_2, and the minimum A required.

***2.27** (a) Use Eq. (2.5) to show that a reduction ΔA in the op-amp gain A gives rise to a reduction $\Delta|G|$ in the magnitude of the closed-loop gain G with $\Delta|G|$ and ΔA related by

$$\frac{\Delta|G|/|G|}{\Delta A/A} \simeq \frac{1 + R_2/R_1}{A}$$

Assume that $\left(1 + \frac{R_2}{R_1}\right) \ll A$ and $\frac{\Delta A}{A} \ll 1$.

(b) If in a closed-loop amplifier with a nominal gain (i.e., R_2/R_1) of 100, A decreases by 10%, what is the minimum nominal A required to limit the percentage change in $|G|$ to 0.1%?

2.28 Consider the circuit in Fig. 2.8 with $R_1 = R_2 = R_4 = 1$ MΩ, and assume the op amp to be ideal. Find values for R_3 to obtain the following gains:

(a) -100 V/V
(b) -10 V/V
(c) -2 V/V

D 2.29 An inverting op-amp circuit using an ideal op amp must be designed to have a gain of -500 V/V using resistors no larger than 100 kΩ.

(a) For the simple two-resistor circuit, what input resistance would result?
(b) If the circuit in Fig. 2.8 is used with three resistors of maximum value, what input resistance results? What is the value of the smallest resistor needed?

2.30 The inverting circuit with the T network in the feedback is redrawn in Fig. P2.30 in a way that emphasizes the observation that R_2 and R_3 in effect are in parallel (because the ideal op amp forces a virtual ground at the inverting input terminal). Use this observation to derive an expression for the gain (v_O/v_I) by first finding (v_X/v_I) and (v_O/v_X). For the latter use the voltage-divider rule applied to R_4 and ($R_2 \parallel R_3$).

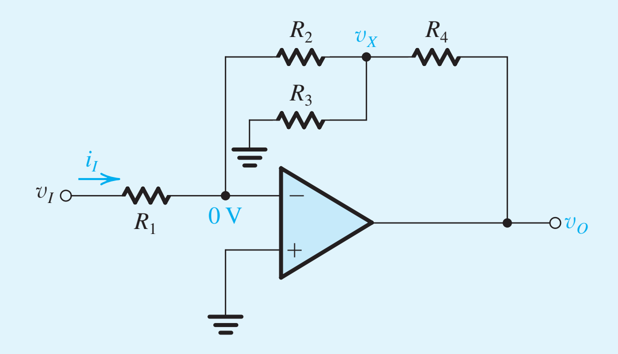


Figure P2.30

***2.31** The circuit in Fig. P2.31 can be considered to be an extension of the circuit in Fig. 2.8.

(a) Find the resistances looking into node 1, R_1; node 2, R_2; node 3, R_3; and node 4, R_4.

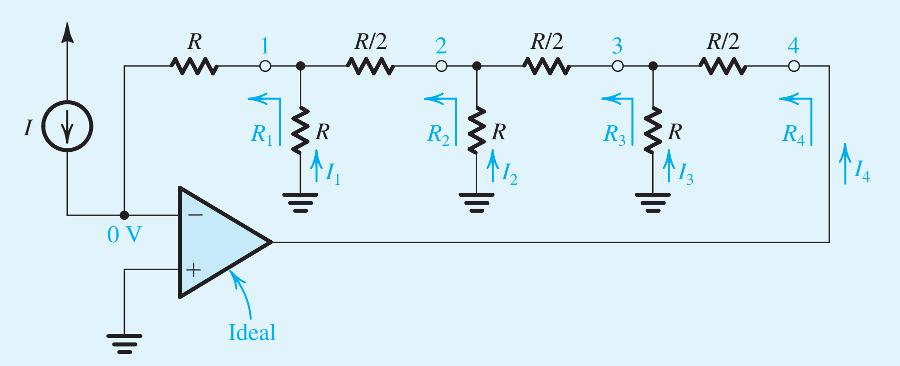


Figure P2.31

(b) Find the currents I_1, I_2, I_3, and I_4, in terms of the input current I.
(c) Find the voltages at nodes 1, 2, 3, and 4, that is, V_1, V_2, V_3, and V_4 in terms of (IR).

2.32 The circuit in Fig. P2.32 utilizes an ideal op amp.

(a) Find I_1, I_2, I_3, I_L, and V_x.
(b) If V_O is not to be lower than −13 V, find the maximum allowed value for R_L.
(c) If R_L is varied in the range 100 Ω to 1 kΩ, what is the corresponding change in I_L and in V_O?

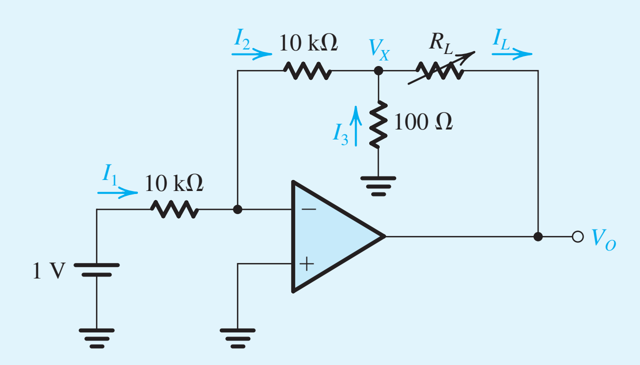


Figure P2.32

D 2.33 Use the circuit in Fig. P2.32 as an inspiration to design a circuit that supplies a constant current I_L of 3.1 mA to a variable resistance R_L. Assume the availability of a 1.5-V battery and design so that the current drawn from the battery is 0.1 mA. For the smallest resistance in the circuit, use 500 Ω. If the op amp saturates at ±10 V, what is the maximum value that R_L can have while the current source supplying it operates properly?

D 2.34 Assuming the op amp to be ideal, it is required to design the circuit shown in Fig. P2.34 to implement a current amplifier with gain $i_L/i_I = 11$ A/A.

(a) Find the required value for R.
(b) What are the input and the output resistance of this current amplifier?
(c) If $R_L = 1$ kΩ and the op amp operates in an ideal manner as long as v_O is in the range ±12 V, what range of i_I is possible?
(d) If the amplifier is fed with a current source having a current of 0.2 mA and a source resistance of 10 kΩ, find i_L.

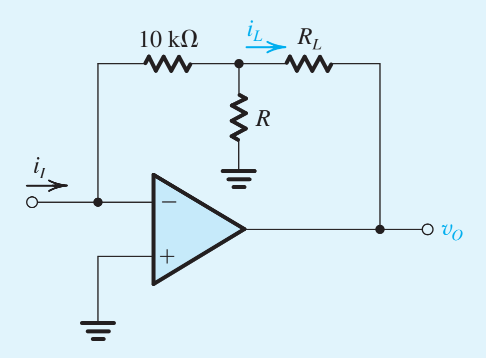


Figure P2.34

D 2.35 Design the circuit shown in Fig. P2.35 to have an input resistance of 100 kΩ and a gain that can be varied from −1 V/V to −100 V/V using the 100-kΩ potentiometer R_4. What voltage gain results when the potentiometer is set exactly at its middle value?

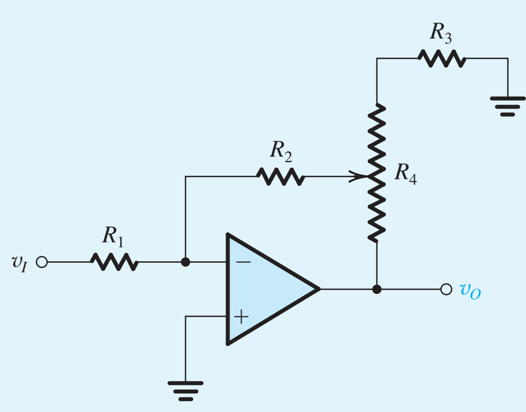


Figure P2.35

2.36 A weighted summer circuit using an ideal op amp has three inputs using 10-kΩ resistors and a feedback resistor of 50 kΩ. A signal v_1 is connected to two of the inputs while a signal v_2 is connected to the third. Express v_O in terms of v_1 and v_2. If $v_1 = 1$ V and $v_2 = -1$ V, what is v_O?

D 2.37 Design an op-amp circuit to provide an output $v_O = -[2v_1 + (v_2/2)]$. Choose relatively low values of resistors but ones for which the input current (from each input signal source) does not exceed 50 μA for 1-V input signals.

D 2.38 Use the scheme illustrated in Fig. 2.10 to design an op-amp circuit with inputs v_1, v_2, and v_3, whose output is

$v_O = -(2v_1 + 4v_2 + 8v_3)$ using small resistors but no smaller than 1 kΩ.

D 2.39 An ideal op amp is connected in the weighted summer configuration of Fig. 2.10. The feedback resistor $R_f = 100$ kΩ, and six 100-kΩ resistors are connected to the inverting input terminal of the op amp. Show, by sketching the various circuit configurations, how this basic circuit can be used to implement the following functions:

(a) $v_O = -(v_1 + 2v_2 + 3v_3)$
(b) $v_O = -(v_1 + v_2 + 2v_3 + 2v_4)$
(c) $v_O = -(v_1 + 5v_2)$
(d) $v_O = -6v_1$

In each case find the input resistance seen by each of the signal sources supplying v_1, v_2, v_3, and v_4. Suggest at least two additional summing functions that you can realize with this circuit. How would you realize a summing coefficient that is 0.5?

D 2.40 Give a circuit, complete with component values, for a weighted summer that shifts the dc level of a sine-wave signal of 3 sin(ωt) V from zero to −3 V. Assume that in addition to the sine-wave signal you have a dc reference voltage of 1.5 V available. Sketch the output signal waveform.

D 2.41 Use two ideal op amps and resistors to implement the summing function

$$v_O = v_1 + 2v_2 - 3v_3 - 5v_4$$

D 2.42 In an instrumentation system, there is a need to take the difference between two signals, one of $v_1 = 2\sin(2\pi \times 60t) + 0.01\sin(2\pi \times 1000t)$ volts and another of $v_2 = 2\sin(2\pi \times 60t) - 0.01\sin(2\pi \times 1000t)$ volts. Draw a circuit that finds the required difference using two op amps and mainly 100-kΩ resistors. Since it is desirable to amplify the 1000-Hz component in the process, arrange to provide an overall gain of 100 as well. The op amps available are ideal except that their output voltage swing is limited to ±10 V.

***2.43** Figure P2.43 shows a circuit for a digital-to-analog converter (DAC). The circuit accepts a 4-bit input binary word $a_3a_2a_1a_0$, where a_0, a_1, a_2, and a_3 take the values of 0 or 1, and it provides an analog output voltage v_O proportional to the value of the digital input. Each of the bits of the input word controls the correspondingly numbered switch. For instance, if a_2 is 0 then switch S_2 connects the 20-kΩ resistor to ground, while if a_2 is 1 then S_2 connects the 20-kΩ resistor to the +5-V power supply. Show that v_O is given by

$$v_O = -\frac{R_f}{16}[2^0 a_0 + 2^1 a_1 + 2^2 a_2 + 2^3 a_3]$$

where R_f is in kilohms. Find the value of R_f so that v_O ranges from 0 to −12 volts.

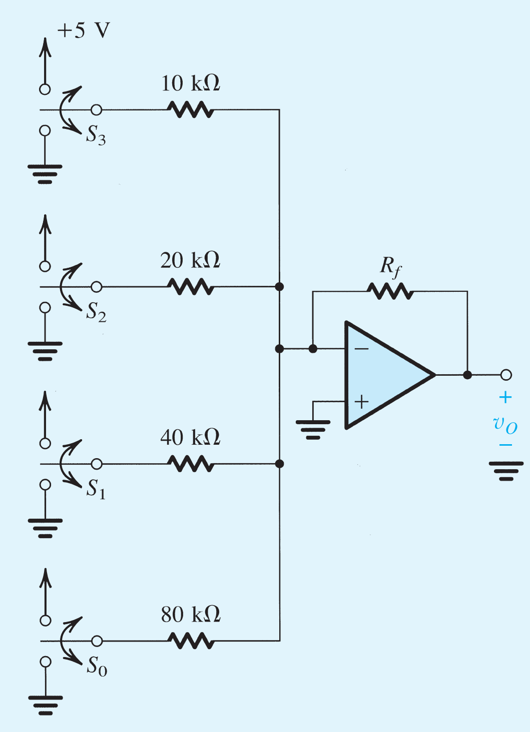


Figure P2.43

Section 2.3: The Noninverting Configuration

D 2.44 Given an ideal op amp to implement designs for the following closed-loop gains, what values of resistors (R_1, R_2) should be used? Where possible, use at least one 10-kΩ resistor as the smallest resistor in your design.

(a) +1 V/V
(b) +2 V/V
(c) +21 V/V
(d) +100 V/V

D 2.45 Design a circuit based on the topology of the noninverting amplifier to obtain a gain of +1.5 V/V, using only 10-kΩ resistors. Note that there are two possibilities. Which of these can be easily converted to have a gain of

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

either +1.0 V/V or +2.0 V/V simply by short-circuiting a single resistor in each case?

D 2.46 Figure P2.46 shows a circuit for an analog voltmeter of very high input resistance that uses an inexpensive moving-coil meter. The voltmeter measures the voltage V applied between the op amp's positive-input terminal and ground. Assuming that the moving coil produces full-scale deflection when the current passing through it is 100 μA, find the value of R such that a full-scale reading is obtained when V is +10 V. Does the meter resistance shown affect the voltmeter calibration?

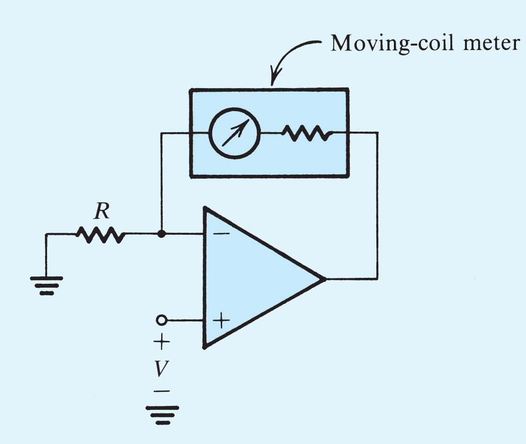


Figure P2.46

D *2.47 (a) Use superposition to show that the output of the circuit in Fig. P2.47 is given by

$$v_O = \left[\frac{R_f}{R_{N1}}v_{N1} + \frac{R_f}{R_{N2}}v_{N2} + \cdots + \frac{R_f}{R_{Nn}}v_{Nn}\right] + \left[1 + \frac{R_f}{R_N}\right]\left[\frac{R_P}{R_{P1}}v_{P1} + \frac{R_P}{R_{P2}}v_{P2} + \cdots + \frac{R_P}{R_{Pn}}v_{Pn}\right]$$

where $R_N = R_{N1} \| R_{N2} \| \cdots \| R_{Nn}$, and

$$R_P = R_{P1} \| R_{P2} \| \cdots \| R_{Pn} \| R_{P0}$$

(b) Design a circuit to obtain

$$v_O = -4v_{N1} + v_{P1} + 3v_{P2}$$

The smallest resistor used should be 10 kΩ.

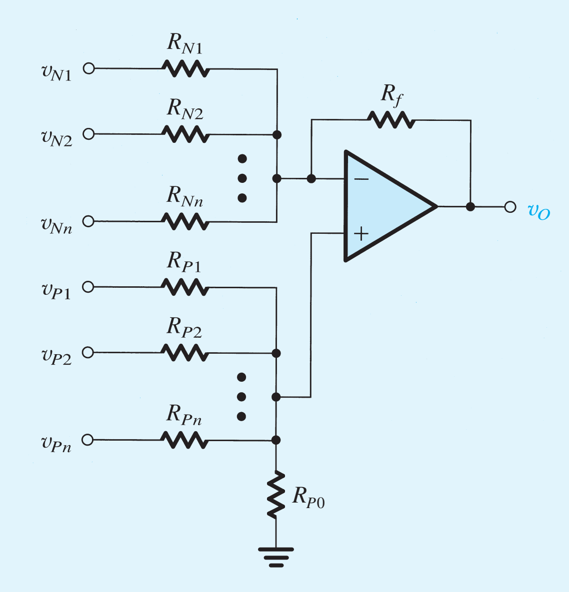


Figure P2.47

D *2.48 Design a circuit, using one ideal op amp, whose output is $v_O = v_{I1} + 2v_{I2} - 9v_{I3} + 4v_{I4}$. (*Hint:* Use a structure similar to that shown in general form in Fig. P2.47.)

2.49 Derive an expression for the voltage gain, v_O/v_I, of the circuit in Fig. P2.49.

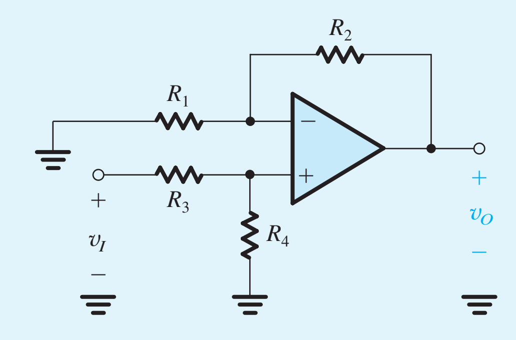


Figure P2.49

2.50 For the circuit in Fig. P2.50, use superposition to find v_O in terms of the input voltages v_1 and v_2. Assume an ideal op amp. For

$$v_1 = 10\sin(2\pi \times 60t) - 0.1\sin(2\pi \times 1000t), \text{ volts}$$
$$v_2 = 10\sin(2\pi \times 60t) + 0.1\sin(2\pi \times 1000t), \text{ volts}$$

find v_O.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

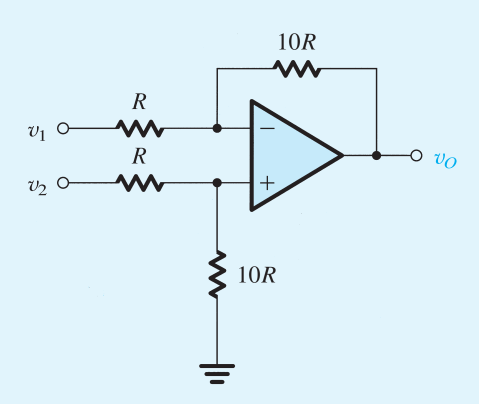


Figure P2.50

D 2.51 The circuit shown in Fig. P2.51 utilizes a 10-kΩ potentiometer to realize an adjustable-gain amplifier. Derive an expression for the gain as a function of the potentiometer setting x. Assume the op amp to be ideal. What is the range of gains obtained? Show how to add a fixed resistor so that the gain range can be 1 to 11 V/V. What should the resistor value be?

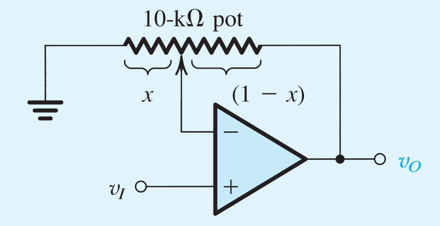


Figure P2.51

D 2.52 Given the availability of resistors of value 1 kΩ and 10 kΩ only, design a circuit based on the noninverting configuration to realize a gain of +10 V/V. What is the input resistance of your amplifier?

2.53 It is required to connect a 10-V source with a source resistance of 1 MΩ to a 1-kΩ load. Find the voltage that will appear across the load if:

(a) The source is connected directly to the load.
(b) A unity-gain op-amp buffer is inserted between the source and the load.

In each case find the load current and the current supplied by the source. Where does the load current come from in case (b)?

2.54 Derive an expression for the gain of the voltage follower of Fig. 2.14, assuming the op amp to be ideal except for having a finite gain A. Calculate the value of the closed-loop gain for $A = 1000$, 100, and 10. In each case find the percentage error in gain magnitude from the nominal value of unity.

2.55 Complete the following table for feedback amplifiers created using one ideal op amp. Note that R_{in} signifies input resistance and R_1 and R_2 are feedback-network resistors as labeled in the inverting and noninverting configurations.

Case	Gain	R_{in}	R_1	R_2
a	−10 V/V	10 kΩ		
b	−1 V/V		100 kΩ	
c	−2 V/V			200 kΩ
d	+1 V/V	∞		
e	+2 V/V		100 kΩ	
f	+11 V/V			100 kΩ
g	−0.5 V/V	20 kΩ		

D 2.56 A noninverting op-amp circuit with nominal gain of 10 V/V uses an op amp with open-loop gain of 100 V/V and a lowest-value resistor of 10 kΩ. What closed-loop gain actually results? With what value resistor can which resistor be shunted to achieve the nominal gain? If in the manufacturing process, an op amp of gain 200 V/V were used, what closed-loop gain would result in each case (the uncompensated one, and the compensated one)?

2.57 Use Eq. (2.11) to show that if the reduction in the closed-loop gain G from the nominal value $G_0 = 1 + R_2/R_1$ is to be kept less than x% of G_0, then the open-loop gain of the op amp must exceed G_0 by at least a factor $F = (100/x) - 1 \simeq 100/x$. Find the required F for $x = 0.01$, 0.1, 1, and 10. Utilize these results to find for each value of x the minimum required open-loop gain to obtain closed-loop gains of 1, 10, 10^2, 10^3, and 10^4 V/V.

2.58 For each of the following combinations of op-amp open-loop gain A and nominal closed-loop gain G_0, calculate the actual closed-loop gain G that is achieved. Also, calculate the percentage by which $|G|$ falls short of the nominal gain magnitude $|G_0|$.

Case	G_0 (V/V)	A (V/V)
a	−1	10
b	+1	10
c	−1	100
d	+10	10
e	−10	100
f	−10	1000
g	+1	2

2.59 Figure P2.59 shows a circuit that provides an output voltage v_O whose value can be varied by turning the wiper of the 100-kΩ potentiometer. Find the range over which v_O can be varied. If the potentiometer is a "20-turn" device, find the change in v_O corresponding to each turn of the pot.

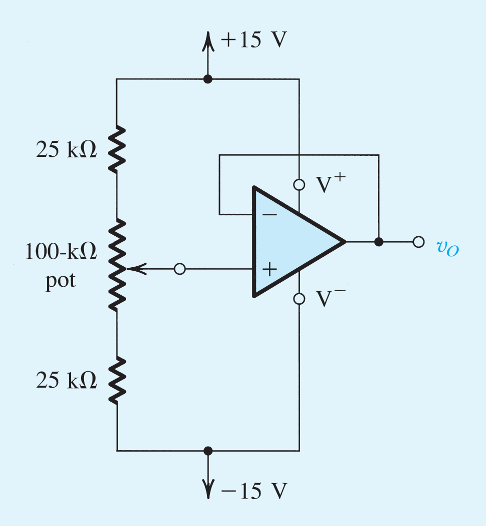


Figure P2.59

Section 2.4: Difference Amplifiers

2.60 Find the voltage gain v_O/v_{Id} for the difference amplifier of Fig. 2.16 for the case $R_1 = R_3 = 5$ kΩ and $R_2 = R_4 = 100$ kΩ. What is the differential input resistance R_{id}? If the two key resistance ratios (R_2/R_1) and (R_4/R_3) are different from each other by 1%, what do you expect the common-mode gain A_{cm} to be? Also, find the CMRR in this case. Neglect the effect of the ratio mismatch on the value of A_d.

D 2.61 Using the difference amplifier configuration of Fig. 2.16 and assuming an ideal op amp, design the circuit to provide the following differential gains. In each case, the differential input resistance should be 20 kΩ.

(a) 1 V/V
(b) 5 V/V
(c) 100 V/V
(d) 0.5 V/V

2.62 For the circuit shown in Fig. P2.62, express v_O as a function of v_1 and v_2. What is the input resistance seen by v_1 alone? By v_2 alone? By a source connected between the two input terminals? By a source connected to both input terminals simultaneously?

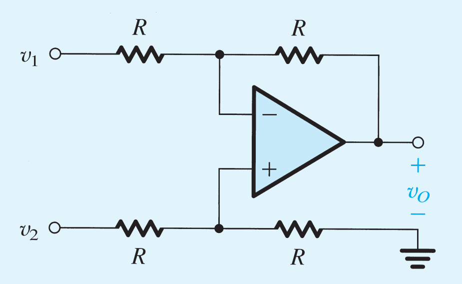


Figure P2.62

2.63 Consider the difference amplifier of Fig. 2.16 with the two input terminals connected together to an input common-mode signal source. For $R_2/R_1 = R_4/R_3$, show that the input common-mode resistance is $(R_3 + R_4) \| (R_1 + R_2)$.

2.64 Consider the circuit of Fig. 2.16, and let each of the v_{I1} and v_{I2} signal sources have a series resistance R_s. What condition must apply in addition to the condition in Eq. (2.15) in order for the amplifier to function as an ideal difference amplifier?

***2.65** For the difference amplifier shown in Fig. P2.62, let all the resistors be 10 kΩ ± x%. Find an expression for the worst-case common-mode gain that results. Evaluate this for $x = 0.1$, 1, and 5. Also, evaluate the resulting CMRR in each case. Neglect the effect of resistor tolerances on A_d.

2.66 For the difference amplifier of Fig. 2.16, show that if each resistor has a tolerance of ±100ϵ% (i.e., for, say, a 5% resistor, $\epsilon = 0.05$) then the worst-case CMRR is given approximately by

$$\text{CMRR} \simeq 20\log\left[\frac{K+1}{4\epsilon}\right]$$

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

where K is the nominal (ideal) value of the ratios (R_2/R_1) and (R_4/R_3). Calculate the value of worst-case CMRR for an amplifier designed to have a differential gain of ideally 100 V/V, assuming that the op amp is ideal and that 1% resistors are used. What resistor tolerance is needed if a CMRR of 80 dB is required?

D *2.67 Design the difference amplifier circuit of Fig. 2.16 to realize a differential gain of 1000, a differential input resistance of 2 kΩ, and a minimum CMRR of 88 dB. Assume the op amp to be ideal. Specify both the resistor values and their required tolerance (e.g., better than x%).

***2.68** (a) Find A_d and A_{cm} for the difference amplifier circuit shown in Fig. P2.68.
(b) If the op amp is specified to operate properly as long as the common-mode voltage at its positive and negative inputs falls in the range ±2.5 V, what is the corresponding limitation on the range of the input common-mode signal v_{Icm}? (This is known as the **common-mode range** of the differential amplifier.)
(c) The circuit is modified by connecting a 10-kΩ resistor between node A and ground, and another 10-kΩ resistor between node B and ground. What will now be the values of A_d, A_{cm}, and the input common-mode range?

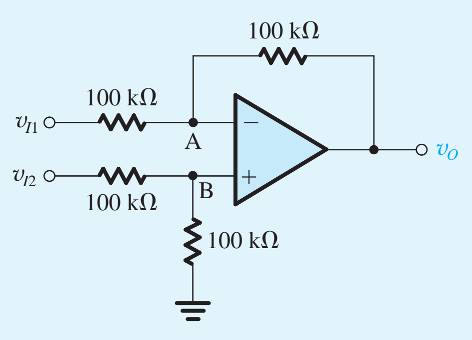


Figure P2.68

D *2.69 To obtain a high-gain, high-input-resistance difference amplifier, the circuit in Fig. P2.69 employs positive feedback, in addition to the negative feedback provided by the resistor R connected from the output to the negative input of the op amp. Specifically, a voltage divider (R_5, R_6) connected across the output feeds a fraction β of the output, that is, a voltage βv_O, back to the positive-input terminal of the op amp through a resistor R. Assume that R_5 and R_6 are much smaller than R so that the current through R is much lower than the current in the voltage divider, with the result that $\beta \simeq R_6/(R_5+R_6)$. Show that the differential gain is given by

$$A_d \equiv \frac{v_O}{v_{Id}} = \frac{1}{1-\beta}$$

(*Hint:* Use superposition.)

Design the circuit to obtain a differential gain of 10 V/V and differential input resistance of 2 MΩ. Select values for R, R_5, and R_6, such that $(R_5+R_6) \leq R/100$.

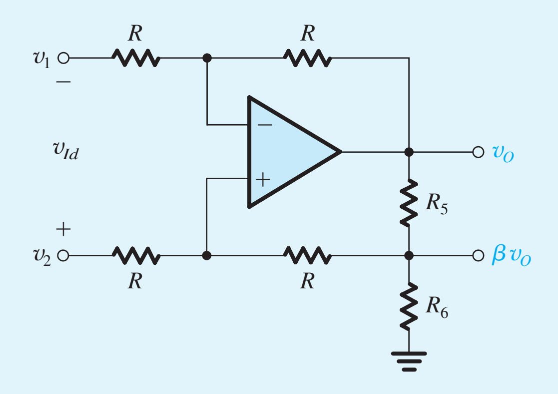


Figure P2.69

***2.70** Figure P2.70 shows a modified version of the difference amplifier. The modified circuit includes a resistor R_G, which can be used to vary the gain. Show that the differential voltage gain is given by

$$\frac{v_O}{v_{Id}} = -2\frac{R_2}{R_1}\left[1+\frac{R_2}{R_G}\right]$$

(*Hint:* The virtual short circuit at the op-amp input causes the current through the R_1 resistors to be $v_{Id}/2R_1$).

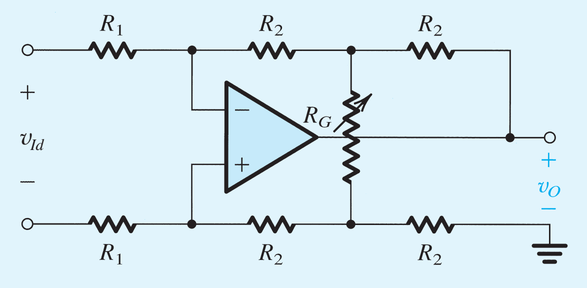


Figure P2.70

D *2.71 The circuit shown in Fig. P2.71 is a representation of a versatile, commercially available IC, the INA105, manufactured by Burr-Brown and known as a differential amplifier module. It consists of an op amp and precision, laser-trimmed, metal-film resistors. The circuit can be configured for a variety of applications by the appropriate connection of terminals A, B, C, D, and O.

(a) Show how the circuit can be used to implement a difference amplifier of unity gain.
(b) Show how the circuit can be used to implement single-ended amplifiers with gains:
 (i) −1 V/V
 (ii) +1 V/V
 (iii) +2 V/V
 (iv) +1/2 V/V

Avoid leaving a terminal open-circuited, for such a terminal may act as an "antenna," picking up interference and noise through capacitive coupling. Rather, find a convenient node to connect such a terminal in a redundant way. When more than one circuit implementation is possible, comment on the relative merits of each, taking into account such considerations as dependence on component matching and input resistance.

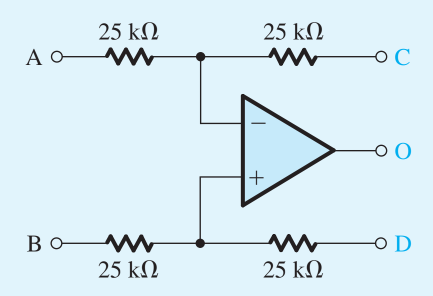


Figure P2.71

2.72 Consider the instrumentation amplifier of Fig. 2.20(b) with a common-mode input voltage of +3 V (dc) and a differential input signal of 100-mV peak sine wave. Let $2R_1 = 2\ \text{k}\Omega$, $R_2 = 50\ \text{k}\Omega$, $R_3 = R_4 = 10\ \text{k}\Omega$. Find the voltage at every node in the circuit.

2.73 (a) Consider the instrumentation amplifier circuit of Fig. 2.20(a). If the op amps are ideal except that their outputs saturate at ±12 V, in the manner shown in Fig. 1.14, find the maximum allowed input common-mode signal for the case $R_1 = 1\ \text{k}\Omega$ and $R_2 = 100\ \text{k}\Omega$.
(b) Repeat (a) for the circuit in Fig. 2.20(b), and comment on the difference between the two circuits.

2.74 (a) Expressing v_{I1} and v_{I2} in terms of differential and common-mode components, find v_{O1} and v_{O2} in the circuit in Fig. 2.20(a) and hence find their differential component $v_{O2} - v_{O1}$ and their common-mode component $\frac{1}{2}(v_{O1} + v_{O2})$. Now find the differential gain and the common-mode gain of the first stage of this instrumentation amplifier and hence the CMRR.
(b) Repeat for the circuit in Fig. 2.20(b), and comment on the difference between the two circuits.

***2.75** For an instrumentation amplifier of the type shown in Fig. 2.20(b), a designer proposes to make $R_2 = R_3 = R_4 = 100\ \text{k}\Omega$, and $2R_1 = 10\ \text{k}\Omega$. For ideal components, what difference-mode gain, common-mode gain, and CMRR result? Reevaluate the worst-case values for these for the situation in which all resistors are specified as ±1% units. Repeat the latter analysis for the case in which $2R_1$ is reduced to 1 kΩ. What do you conclude about the effect of the gain of the first stage on CMRR? (*Hint:* Eq. (2.19) can be used to evaluate A_{cm} of the second stage.)

D 2.76 Design the instrumentation-amplifier circuit of Fig. 2.20(b) to realize a differential gain, variable in the range 2 to 100, utilizing a 100-kΩ pot as variable resistor.

SIM *2.77 The circuit shown in Fig. P2.77 is intended to supply a voltage to floating loads (those for which both terminals are ungrounded) while making greatest possible use of the available power supply.

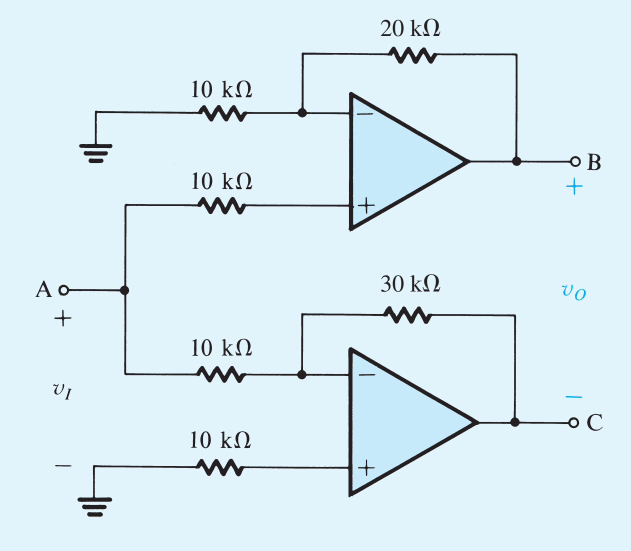


Figure P2.77

(a) Assuming ideal op amps, sketch the voltage waveforms at nodes B and C for a 1-V peak-to-peak sine wave applied at A. Also sketch v_O.
(b) What is the voltage gain v_O/v_I?
(c) Assuming that the op amps operate from ±15-V power supplies and that their output saturates at ±14 V (in the manner shown in Fig. 1.14), what is the largest sine-wave output that can be accommodated? Specify both its peak-to-peak and rms values.

***2.78** The two circuits in Fig. P2.78 are intended to function as voltage-to-current converters; that is, they supply the load impedance Z_L with a current proportional to v_I and independent of the value of Z_L. Show that this is indeed the case, and find for each circuit i_O as a function of v_I. Comment on the differences between the two circuits.

Section 2.5: Integrators and Differentiators

2.79 A Miller integrator incorporates an ideal op amp, a resistor R of 10 kΩ, and a capacitor C of 1 nF. A sine-wave signal is applied to its input.

(a) At what frequency (in Hz) are the input and output signals equal in amplitude?
(b) At that frequency, how does the phase of the output sine wave relate to that of the input?
(c) If the frequency is lowered by a factor of 10 from that found in (a), by what factor does the output voltage change, and in what direction (smaller or larger)?
(d) What is the phase relation between the input and output in situation (c)?

D 2.80 Design a Miller integrator with a time constant of 1 s and an input resistance of 100 kΩ. A dc voltage of −1 volt is applied at the input at time 0, at which moment $v_O = -10$ V. How long does it take the output to reach 0 V? +10 V?

2.81 An op-amp-based inverting integrator is measured at 10 kHz to have a voltage gain of −100 V/V. At what frequency is its gain reduced to −1 V/V? What is the integrator time constant?

D 2.82 Design a Miller integrator that has a unity-gain frequency of 10 krad/s and an input resistance of 100 kΩ. Sketch the output you would expect for the situation in which, with output initially at 0 V, a 2-V, 100-µs pulse is applied to the input. Characterize the output that results when a sine wave 2 sin $10^4 t$ is applied to the input.

D 2.83 Design a Miller integrator whose input resistance is 10 kΩ and unity-gain frequency is 100 kHz. What components are needed? For long-term stability, a feedback resistor is introduced across the capacitor to limit the dc gain

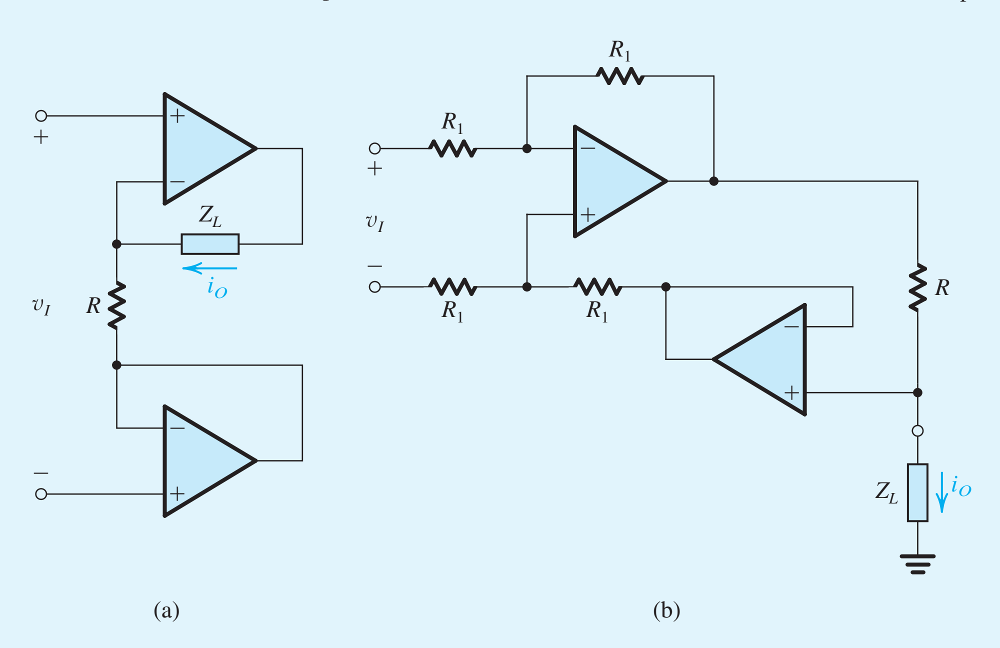


Figure P2.78

to 40 dB. What is its value? What is the associated lower 3-dB frequency? Sketch and label the output that results with a 10-μs, 1-V positive-input pulse (initially at 0 V) with (a) no dc stabilization (but with the output initially at 0 V) and (b) the feedback resistor connected.

***2.84** A Miller integrator whose input and output voltages are initially zero and whose time constant is 1 ms is driven by the signal shown in Fig. P2.84. Sketch and label the output waveform that results. Indicate what happens if the input levels are ±2 V, with the time constant the same (1 ms) and with the time constant raised to 2 ms.

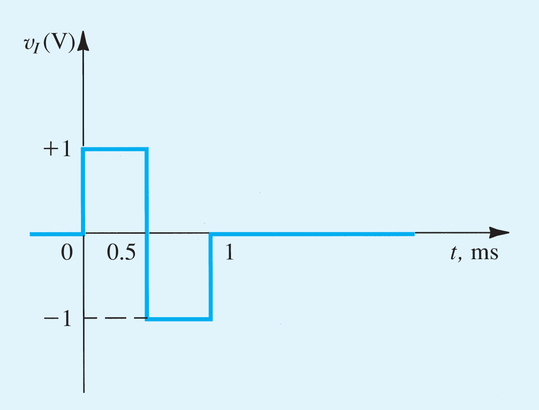


Figure P2.84

2.85 Consider a Miller integrator having a time constant of 1 ms and an output that is initially zero, when fed with a string of pulses of 10-μs duration and 1-V amplitude rising from 0 V (see Fig. P2.85). Sketch and label the output waveform resulting. How many pulses are required for an output voltage change of 1 V?

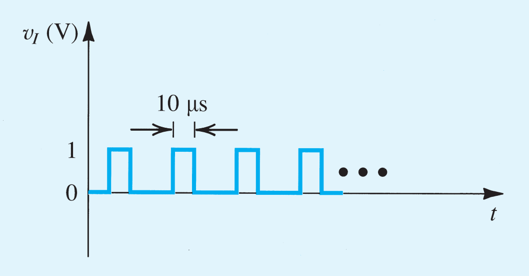


Figure P2.85

D 2.86 Figure P2.86 shows a circuit that performs a low-pass STC function. Such a circuit is known as a first-order, low-pass active filter. Derive the transfer function and show that the dc gain is $(-R_2/R_1)$ and the 3-dB frequency $\omega_0 = 1/CR_2$. Design the circuit to obtain an input resistance of 10 kΩ, a dc gain of 40 dB, and a 3-dB frequency of 1 kHz. At what frequency does the magnitude of the transfer function reduce to unity?

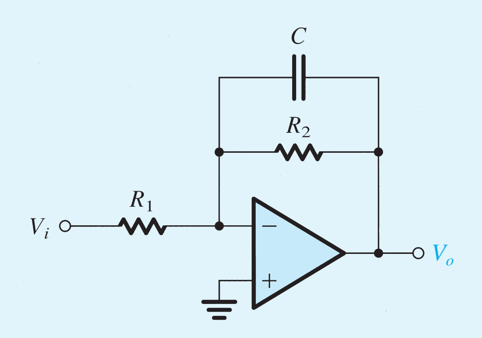


Figure P2.86

***2.87** Show that a Miller integrator implemented with an op amp with open-loop gain A_0 has a low-pass STC transfer function. What is the pole frequency of the STC function? How does this compare with the pole frequency of the ideal integrator? If an ideal Miller integrator is fed with a −1-V pulse signal with a width $T = CR$, what will the output voltage be at $t = T$? Assume that at $t = 0, v_O = 0$. Repeat for an integrator with an op amp having $A_0 = 1000$.

2.88 A differentiator utilizes an ideal op amp, a 10-kΩ resistor, and a 1-nF capacitor. What is the frequency f_0 (in Hz) at which its input and output sine-wave signals have equal magnitude? What is the output signal for a 1-V peak-to-peak sine-wave input with frequency equal to $10f_0$?

2.89 An op-amp differentiator with 1-ms time constant is driven by the rate-controlled step shown in Fig. P2.89. Assuming v_O to be zero initially, sketch and label its waveform.

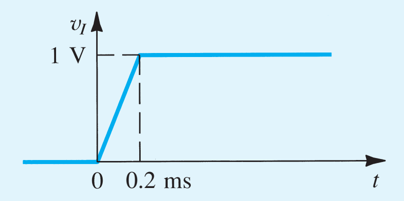


Figure P2.89

2.90 An op-amp differentiator, employing the circuit shown in Fig. 2.27(a), has $R = 20\ \text{k}\Omega$ and $C = 0.1\ \mu\text{F}$. When a triangle wave of ± 1-V peak amplitude at 1 kHz is applied to the input, what form of output results? What is its frequency? What is its peak amplitude? What is its average value? What value of R is needed to cause the output to have a 12-V peak amplitude?

2.91 Use an ideal op amp to design a differentiation circuit for which the time constant is 10^{-3} s using a 10-nF capacitor. What are the gains and phase shifts found for this circuit at one-tenth and 10 times the unity-gain frequency? A series input resistor is added to limit the gain magnitude at high frequencies to 100 V/V. What is the associated 3-dB frequency? What gain and phase shift result at 10 times the unity-gain frequency?

D 2.92 Figure P2.92 shows a circuit that performs the high-pass, single-time-constant function. Such a circuit is known as a first-order high-pass active filter. Derive the transfer function and show that the high-frequency gain is $(-R_2/R_1)$ and the 3-dB frequency $\omega_0 = 1/CR_1$. Design the circuit to obtain a high-frequency input resistance of 1 kΩ, a high-frequency gain of 40 dB, and a 3-dB frequency of 2 kHz. At what frequency does the magnitude of the transfer function reduce to unity?

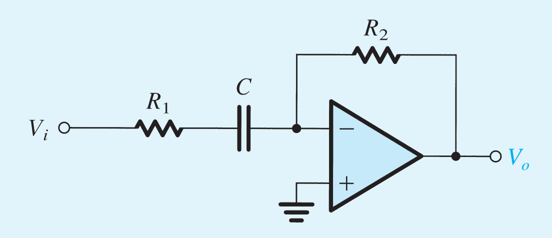


Figure P2.92

D **2.93 Derive the transfer function of the circuit in Fig. P2.93 (for an ideal op amp) and show that it can be written in the form

$$\frac{V_o}{V_i} = \frac{-R_2/R_1}{[1 + (\omega_1/j\omega)][1 + j(\omega/\omega_2)]}$$

where $\omega_1 = 1/C_1R_1$ and $\omega_2 = 1/C_2R_2$. Assuming that the circuit is designed such that $\omega_2 \gg \omega_1$, find approximate expressions for the transfer function in the following frequency regions:

(a) $\omega \ll \omega_1$
(b) $\omega_1 \ll \omega \ll \omega_2$
(c) $\omega \gg \omega_2$

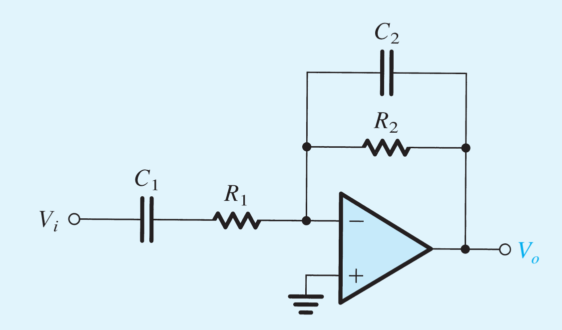


Figure P2.93

Use these approximations to sketch a Bode plot for the magnitude response. Observe that the circuit performs as an amplifier whose gain rolls off at the low-frequency end in the manner of a high-pass STC network, and at the high-frequency end in the manner of a low-pass STC network. Design the circuit to provide a gain of 40 dB in the "middle-frequency range," a low-frequency 3-dB point at 200 Hz, a high-frequency 3-dB point at 200 kHz, and an input resistance (at $\omega \gg \omega_1$) of 2 kΩ.

Section 2.6: DC Imperfections

2.94 An op amp wired in the inverting configuration with the input grounded, having $R_2 = 100\ \text{k}\Omega$ and $R_1 = 2\ \text{k}\Omega$, has an output dc voltage of −0.2 V. If the input bias current is known to be very small, find the input offset voltage.

2.95 A noninverting amplifier with a gain of 100 uses an op amp having an input offset voltage of ±2 mV. Find the output when the input is 0.01 sin ωt, volts.

2.96 A noninverting amplifier with a closed-loop gain of 1000 is designed using an op amp having an input offset voltage of 3 mV and output saturation levels of ±12 V. What is the maximum amplitude of the sine wave that can be applied at the input without the output clipping? If the amplifier is

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

capacitively coupled in the manner indicated in Fig. 2.36, what would the maximum possible amplitude be?

2.97 An op amp connected in a closed-loop inverting configuration having a gain of 1000 V/V and using relatively small-valued resistors is measured with input grounded to have a dc output voltage of −1.8 V. What is its input offset voltage? Prepare an offset-voltage-source sketch resembling that in Fig. 2.28. Be careful of polarities.

2.98 A particular inverting amplifier with nominal gain of −100 V/V uses an imperfect op amp in conjunction with 100-kΩ and 10-MΩ resistors. The output voltage is found to be +5.3 V when measured with the input open and +5 V with the input grounded.

(a) What is the bias current of this amplifier? In what direction does it flow?
(b) Estimate the value of the input offset voltage.
(c) A 10-MΩ resistor is connected between the positive-input terminal and ground. With the input left floating (disconnected), the output dc voltage is measured to be −0.6 V. Estimate the input offset current.

D *2.99 A noninverting amplifier with a gain of +10 V/V using 100 kΩ as the feedback resistor operates from a 5-kΩ source. For an amplifier offset voltage of 0 mV, but with a bias current of 2 μA and an offset current of 0.2 μA, what range of outputs would you expect? Indicate where you would add an additional resistor to compensate for the bias currents. What does the range of possible outputs then become? A designer wishes to use this amplifier with a 15-kΩ source. In order to compensate for the bias current in this case, what resistor would you use? And where?

D 2.100 The circuit of Fig. 2.36 is used to create an ac-coupled noninverting amplifier with a gain of 100 V/V using resistors no larger than 100 kΩ. What values of R_1, R_2, and R_3 should be used? For a break frequency due to C_1 at 100 Hz, and that due to C_2 at 10 Hz, what values of C_1 and C_2 are needed?

***2.101** Consider the difference amplifier circuit in Fig. 2.16. Let $R_1 = R_3 = 10$ kΩ and $R_2 = R_4 = 1$ MΩ. If the op amp has $V_{OS} = 5$ mV, $I_B = 1$ μA, and $I_{OS} = 0.2$ μA, find the worst-case (largest) dc offset voltage at the output.

***2.102** The circuit shown in Fig. P2.102 uses an op amp having a ±3-mV offset. What is its output offset voltage? What does the output offset become with the input ac coupled through a capacitor C? If, instead, a large capacitor is placed in series with a 1-kΩ resistor, what does the output offset become?

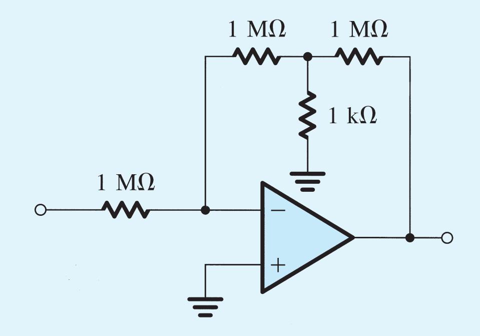


Figure P2.102

2.103 Using offset-nulling facilities provided for the op amp, a closed-loop amplifier with gain of +1000 is adjusted at 25°C to produce zero output with the input grounded. If the input offset-voltage drift is specified to be 20 μV/°C, what output would you expect at 0°C and at 100°C? While nothing can be said separately about the polarity of the output offset at either 0 or 75°C, what would you expect their relative polarities to be?

2.104 An op amp is connected in a closed loop with gain of +100 utilizing a feedback resistor of 1 MΩ.

(a) If the input bias current is 200 nA, what output voltage results with the input grounded?
(b) If the input offset voltage is ±2 mV and the input bias current as in (a), what is the largest possible output that can be observed with the input grounded?
(c) If bias-current compensation is used, what is the value of the required resistor? If the offset current is no more than one-tenth the bias current, what is the resulting output offset voltage (due to offset current alone)?
(d) With bias-current compensation as in (c) in place, what is the largest dc voltage at the output due to the combined effect of offset voltage and offset current?

***2.105** An op amp intended for operation with a closed-loop gain of –100 V/V uses resistors of 10 kΩ and 1 MΩ with a bias-current-compensation resistor R_3. What should the value of R_3 be? With input grounded, the output offset voltage is found to be +0.30 V. Estimate the input offset current assuming zero input offset voltage. If the input offset voltage

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

can be as large as 1 mV of unknown polarity, what range of offset current is possible?

2.106 A Miller integrator with $R = 10\ \text{k}\Omega$ and $C = 10$ nF is implemented by using an op amp with $V_{OS} = 2$ mV, $I_B = 0.1\ \mu$A, and $I_{OS} = 20$ nA. To provide a finite dc gain, a 1-MΩ resistor is connected across the capacitor.

(a) To compensate for the effect of I_B, a resistor is connected in series with the positive-input terminal of the op amp. What should its value be?
(b) With the resistor of (a) in place, find the worst-case dc output voltage of the integrator when the input is grounded.

Section 2.7: Effect of Finite Open-Loop Gain and Bandwidth on Circuit Performance

2.107 The data in the following table apply to internally compensated op amps. Fill in the blank entries.

A_0	f_b (Hz)	f_t (Hz)
10^5	10^2	
10^6		10^6
	10^3	10^8
	10^{-1}	10^6
2×10^5	10	

2.108 A measurement of the open-loop gain of an internally compensated op amp at very low frequencies shows it to be 98 dB; at 100 kHz, this shows it is 40 dB. Estimate values for A_0, f_b, and f_t.

2.109 Measurements of the open-loop gain of a compensated op amp intended for high-frequency operation indicate that the gain is 4×10^3 at 100 kHz and 20×10^3 at 10 kHz. Estimate its 3-dB frequency, its unity-gain frequency, and its dc gain.

2.110 Measurements made on the internally compensated amplifiers listed below provide the dc gain and the frequency at which the gain has dropped by 20 dB. For each, what are the 3 dB and unity-gain frequencies?

(a) 2×10^5 V/V and 5×10^2 Hz
(b) 20×10^5 V/V and 10 Hz
(c) 1800 V/V and 0.1 MHz
(d) 100 V/V and 0.1 GHz
(e) 25 V/mV and 250 kHz

2.111 An inverting amplifier with nominal gain of −50 V/V employs an op amp having a dc gain of 10^4 and a unity-gain frequency of 10^6 Hz. What is the 3-dB frequency $f_{3\text{dB}}$ of the closed-loop amplifier? What is its gain at $0.1 f_{3\text{dB}}$ and at $10 f_{3\text{dB}}$?

2.112 A particular op amp, characterized by a gain–bandwidth product of 20 MHz, is operated with a closed-loop gain of +100 V/V. What 3-dB bandwidth results? At what frequency does the closed-loop amplifier exhibit a −6° phase shift? A −84° phase shift?

2.113 Find the f_t required for internally compensated op amps to be used in the implementation of closed-loop amplifiers with the following nominal dc gains and 3-dB bandwidths:

(a) −50 V/V; 100 kHz
(b) +50 V/V; 100 kHz
(c) +2 V/V; 5 MHz
(d) −2 V/V; 5 MHz
(e) −1000 V/V; 10 kHz
(f) +1 V/V; 1 MHz
(g) −1 V/V; 1 MHz

2.114 A noninverting op-amp circuit with a gain of 96 V/V is found to have a 3-dB frequency of 8 kHz. For a particular system application, a bandwidth of 32 kHz is required. What is the highest gain available under these conditions?

2.115 Consider a unity-gain follower utilizing an internally compensated op amp with $f_t = 2$ MHz. What is the 3-dB frequency of the follower? At what frequency is the gain of the follower 1% below its low-frequency magnitude? If the input to the follower is a 1-V step, find the 10% to 90% rise time of the output voltage. (*Note:* The step response of STC low-pass networks is discussed in Appendix E. Specifically, note that the 10%–90% rise time of a low-pass STC circuit with a time constant τ is 2.2τ.)

D *2.116 It is required to design a noninverting amplifier with a dc gain of 10. When a step voltage of 100 mV is applied at the input, it is required that the output be within 1% of its final value of 1 V in at most 200 ns. What must the f_t of the op amp be? (*Note:* The step response of STC low-pass networks is discussed in Appendix E.)

D *2.117 This problem illustrates the use of cascaded closed-loop amplifiers to obtain an overall bandwidth greater than can be achieved using a single-stage amplifier with the same overall gain.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

(a) Show that cascading two identical amplifier stages, each having a low-pass STC frequency response with a 3-dB frequency f_1, results in an overall amplifier with a 3-dB frequency given by

$$f_{3dB} = \sqrt{\sqrt{2}-1}f_1$$

(b) It is required to design a noninverting amplifier with a dc gain of 40 dB utilizing a single internally compensated op amp with $f_t = 2$ MHz. What is the 3-dB frequency obtained?

(c) Redesign the amplifier of (b) by cascading two identical noninverting amplifiers each with a dc gain of 20 dB. What is the 3-dB frequency of the overall amplifier? Compare this to the value obtained in (b) above.

D **2.118 A designer, wanting to achieve a stable gain of 100 V/V at 5 MHz, considers her choice of amplifier topologies. What unity-gain frequency would a single operational amplifier require to satisfy her need? Unfortunately, the best available amplifier has an f_t of 40 MHz. How many such amplifiers connected in a cascade of identical noninverting stages would she need to achieve her goal? What is the 3-dB frequency of each stage she can use? What is the overall 3-dB frequency?

2.119 Consider the use of an op amp with a unity-gain frequency f_t in the realization of:

(a) An inverting amplifier with dc gain of magnitude K.
(b) A noninverting amplifier with a dc gain of K.

In each case find the 3-dB frequency and the gain–bandwidth product (GBP $\equiv$ |Gain| $\times f_{3dB}$). Comment on the results.

***2.120** Consider an inverting summer with two inputs V_1 and V_2 and with $V_o = -(V_1 + 3V_2)$. Find the 3-dB frequency of each of the gain functions V_o/V_1 and V_o/V_2 in terms of the op amp f_t. (*Hint:* In each case, the other input to the summer can be set to zero—an application of superposition.)

Section 2.8: Large-Signal Operation of Op Amps

2.121 A particular op amp using ±15-V supplies operates linearly for outputs in the range −14 V to +14 V. If used in an inverting amplifier configuration of gain −100, what is the rms value of the largest possible sine wave that can be applied at the input without output clipping?

2.122 Consider an op amp connected in the inverting configuration to realize a closed-loop gain of −100 V/V utilizing resistors of 1 kΩ and 100 kΩ. A load resistance R_L is connected from the output to ground, and a low-frequency sine-wave signal of peak amplitude V_p is applied to the input. Let the op amp be ideal except that its output voltage saturates at ±10 V and its output current is limited to the range ±20 mA.

(a) For $R_L = 1$ kΩ, what is the maximum possible value of V_p while an undistorted output sinusoid is obtained?
(b) Repeat (a) for $R_L = 200$ Ω.
(c) If it is desired to obtain an output sinusoid of 10-V peak amplitude, what minimum value of R_L is allowed?

2.123 An op amp having a slew rate of 10 V/μs is to be used in the unity-gain follower configuration, with input pulses that rise from 0 to 2 V. What is the shortest pulse that can be used while ensuring full-amplitude output? For such a pulse, describe the output resulting.

2.124 For operation with 10-V output pulses with the requirement that the sum of the rise and fall times represent only 20% of the pulse width (at half-amplitude), what is the slew-rate requirement for an op amp to handle pulses 2 μs wide? (*Note:* The rise and fall times of a pulse signal are usually measured between the 10%- and 90%-height points.)

2.125 What is the highest frequency of a triangle wave of 10-V peak-to-peak amplitude that can be reproduced by an op amp whose slew rate is 20 V/μs? For a sine wave of the same frequency, what is the maximum amplitude of output signal that remains undistorted?

2.126 For an amplifier having a slew rate of 40 V/μs, what is the highest frequency at which a 20-V peak-to-peak sine wave can be produced at the output?

D *2.127 In designing with op amps one has to check the limitations on the voltage and frequency ranges of operation of the closed-loop amplifier, imposed by the op-amp finite bandwidth (f_t), slew rate (SR), and output saturation ($V_{o\max}$). This problem illustrates the point by considering the use of an op amp with $f_t = 20$ MHz, SR = 10 V/μs, and $V_{o\max} = 10$ V in the design of a noninverting amplifier with a nominal gain of 10. Assume a sine-wave input with peak amplitude V_i.

(a) If $V_i = 0.5$ V, what is the maximum frequency before the output distorts?
(b) If $f = 200$ kHz, what is the maximum value of V_i before the output distorts?
(c) If $V_i = 50$ mV, what is the useful frequency range of operation?
(d) If $f = 50$ kHz, what is the useful input voltage range?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

CHAPTER 3

Semiconductors

IN THIS CHAPTER YOU WILL LEARN

1. The basic properties of semiconductors and in particular silicon, which is the material used to make most of today's electronic circuits.
2. How doping a pure silicon crystal dramatically changes its electrical conductivity, which is the fundamental idea underlying the use of semiconductors in the implementation of electronic devices.
3. The two mechanisms by which current flows in semiconductors: drift and diffusion of charge carriers.
4. The structure and operation of the *pn* junction; a basic semiconductor structure that implements the diode and plays a dominant role in transistors.

Introduction

Thus far we have dealt with electronic circuits, and notably amplifiers, as system building blocks. For instance, in Chapter 2 we learned how to use op amps to design interesting and useful circuits, taking advantage of the terminal characteristics of the op amp and without any knowledge of what is inside the op-amp package. Though interesting and motivating, this approach has its limitations. Indeed, to achieve our goal of preparing the reader to become a proficient circuit designer, we have to go beyond this black-box or system-level abstraction and learn about the basic devices from which electronic circuits are assembled, namely, diodes (Chapter 4) and transistors (Chapters 5 and 6). These solid-state devices are made using semiconductor materials, predominantly silicon.

In this chapter, we briefly introduce the properties and physics of semiconductors. The objective is to provide a basis for understanding the physical operation of diodes and transistors in order to enable their effective use in the design of circuits. Although many of the concepts studied in this chapter apply to semiconductor materials in general, our treatment is heavily biased toward silicon, simply because it is the material used in the vast majority of microelectronic circuits. To complement the material presented here, Appendix A provides a description of the integrated-circuit fabrication process. As discussed in Appendix A, whether our circuit consists of a single transistor or is an **integrated circuit** containing more than 2 billion transistors, it is fabricated in a single silicon crystal, which gives rise to the name **monolithic circuit**. This chapter therefore begins with a study of the crystal structure of semiconductors and introduces the two types of charge carriers available for current conduction: electrons and holes. The most significant property of semiconductors is that their conductivity can be varied over a very wide range through the introduction of

controlled amounts of impurity atoms into the semiconductor crystal in a process called **doping**. Doped semiconductors are discussed in Section 3.2. This is followed by the study in Section 3.3 of the two mechanisms for current flow in semiconductors, namely, carrier drift and carrier diffusion.

Armed with these basic semiconductor concepts, we spend the remainder of the chapter on the study of an important semiconductor structure: the *pn* junction. In addition to being essentially a diode, the *pn* junction is the basic element of the bipolar junction transistor (BJT, Chapter 6) and plays an important role in the operation of field-effect transistors (FETs, Chapter 5).

3.1 Intrinsic Semiconductors

As their name implies, semiconductors are materials whose conductivity lies between that of conductors, such as copper, and insulators, such as glass. There are two kinds of semiconductors: single-element semiconductors, such as germanium and silicon, which are in group IV in the periodic table; and compound semiconductors, such as gallium-arsenide, which are formed by combining elements from groups III and V or groups II and VI. Compound semiconductors are useful in special electronic circuit applications as well as in applications that involve light, such as light-emitting diodes (LEDs). Of the two elemental semiconductors, germanium was used in the fabrication of very early transistors (late 1940s, early 1950s). It was quickly supplanted, however, with silicon, on which today's integrated-circuit technology is almost entirely based. For this reason, we will deal mostly with silicon devices throughout this book.[1]

A silicon atom has four valence electrons, and thus it requires another four to complete its outermost shell. This is achieved by sharing one of its valence electrons with each of its four neighboring atoms. Each pair of shared electrons forms a **covalent bond**. The result is that a crystal of pure or intrinsic silicon has a regular lattice structure, where the atoms are held in their position by the covalent bonds. Figure 3.1 shows a two-dimensional representation of such a structure.

At sufficiently low temperatures, approaching absolute zero (0 K), all the covalent bonds are intact and no electrons are available to conduct electric current. Thus, at such low temperatures, the intrinsic silicon crystal behaves as an insulator.

At room temperature, sufficient thermal energy exists to break some of the covalent bonds, a process known as thermal generation. As shown in Fig. 3.2, when a covalent bond is broken, an electron is freed. The **free electron** can wander away from its parent atom, and it becomes available to conduct electric current if an electric field is applied to the crystal. As the electron leaves its parent atom, it leaves behind a net positive charge, equal to the magnitude of the electron charge. Thus, an electron from a neighboring atom may be attracted to this positive charge, and leaves its parent atom. This action fills up the "hole" that existed in the ionized atom but creates a new hole in the other atom. This process may repeat itself, with the result that we effectively have a positively charged carrier, or **hole**, moving through the silicon crystal structure and being available to conduct electric current. The charge of a hole is equal in magnitude to the charge of an electron. We can thus see that as temperature increases, more covalent bonds are broken and electron–hole pairs are generated. The increase in the numbers of free electrons and holes results in an increase in the conductivity of silicon.

[1]An exception is the subject of gallium arsenide (GaAs) circuits, which though not covered in this edition of the book, is studied in some detail in material provided on the text website.

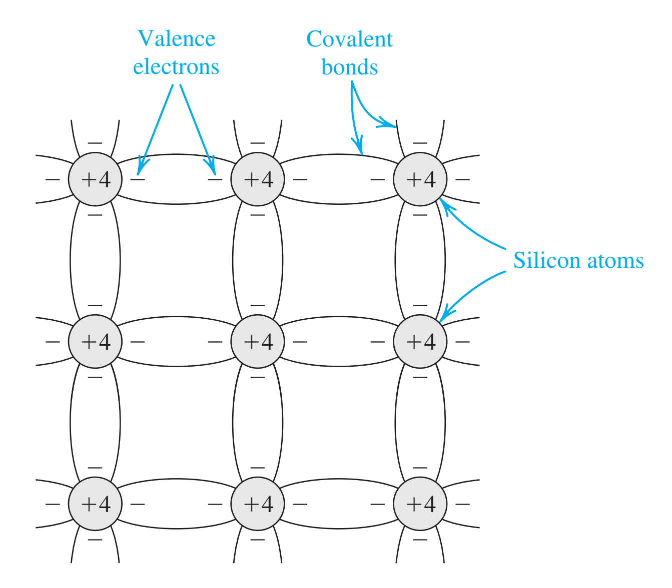


Figure 3.1 Two-dimensional representation of the silicon crystal. The circles represent the inner core of silicon atoms, with +4 indicating its positive charge of $+4q$, which is neutralized by the charge of the four valence electrons. Observe how the covalent bonds are formed by sharing of the valence electrons. At 0 K, all bonds are intact and no free electrons are available for current conduction.

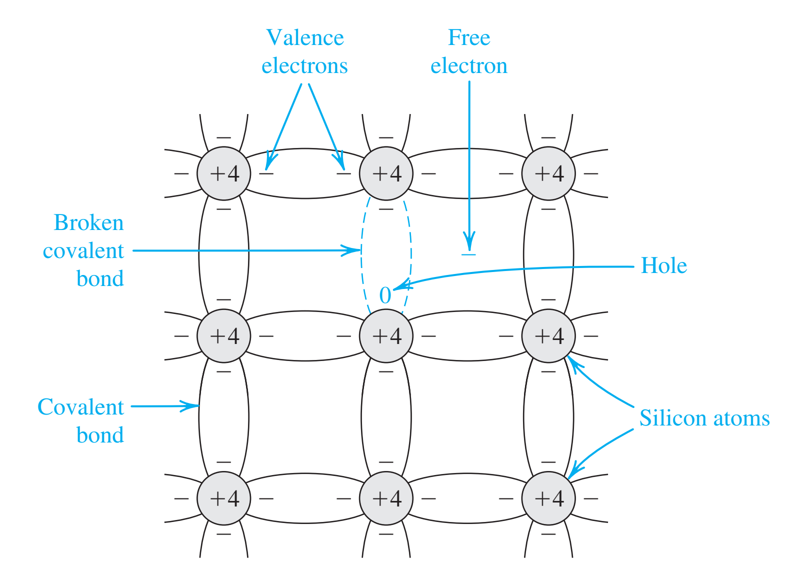


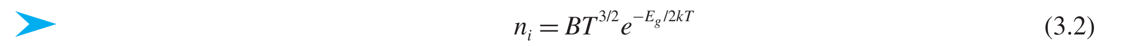
Figure 3.2 At room temperature, some of the covalent bonds are broken by thermal generation. Each broken bond gives rise to a free electron and a hole, both of which become available for current conduction.

Thermal generation results in free electrons and holes in equal numbers and hence equal concentrations, where concentration refers to the number of charge carriers per unit volume (cm^3). The free electrons and holes move randomly through the silicon crystal structure, and in the process some electrons may fill some of the holes. This process, called **recombination**, results in the disappearance of free electrons and holes. The recombination rate is

proportional to the number of free electrons and holes, which in turn is determined by the thermal **generation** rate. The latter is a strong function of temperature. In thermal equilibrium, the recombination rate is equal to the generation rate, and one can conclude that the concentration of free electrons n is equal to the concentration of holes p,

$$n = p = n_i \tag{3.1}$$

where n_i denotes the number of free electrons and holes in a unit volume (cm^3) of intrinsic silicon at a given temperature. Results from semiconductor physics gives n_i as



$$n_i = BT^{3/2}e^{-E_g/2kT} \tag{3.2}$$

where B is a material-dependent parameter that is $7.3 \times 10^{15} cm^{-3}K^{-3/2}$ for silicon; T is the temperature in K; E_g, a parameter known as the **bandgap energy**, is 1.12 electron volt (eV) for silicon[2]; and k is Boltzmann's constant (8.62×10^{-5} eV/K). It is interesting to know that the bandgap energy E_g is the minimum energy required to break a covalent bond and thus generate an electron-hole pair.

Example 3.1

Calculate the value of n_i for silicon at room temperature ($T \simeq 300$ K).

Solution

Substituting the values given above in Eq. (3.2) provides

$$\begin{aligned} n_i &= 7.3 \times 10^{15}(300)^{3/2}e^{-1.12/\left(2 \times 8.62 \times 10^{-5} \times 300\right)} \\ &= 1.5 \times 10^{10} \text{carriers/cm}^3 \end{aligned}$$

Although this number seems large, to place it into context note that silicon has 5×10^{22} atoms/cm^3. Thus at room temperature only one in about 5×10^{12} atoms is ionized and contributing a free electron and a hole!

Finally, it is useful for future purposes to express the product of the hole and free-electron concentration as

$$pn = n_i^2 \tag{3.3}$$

where for silicon at room temperature, $n_i \simeq 1.5 \times 10^{10}/cm^3$. As will be seen shortly, this relationship extends to extrinsic or doped silicon as well.

[2]Note that 1 eV $= 1.6 \times 10^{-19}$ J.

LCDs, THE FACE OF ELECTRONICS:

The existence of liquid crystals whose color could be changed by means of an external heat source was first reported in 1888 by an Austrian botanical physiologist. The LC idea lay dormant until the late 1940s, however. Subsequent developments in the field of solid-state electronics provided the technology to harness the technique in display media, with the first LCDs being demonstrated by RCA beginning in 1962. Today, LCDs are an essential component in every mobile device as the interface to the world of electronics within. At the other end of the scale, large LCDs are used in flat-panel TVs, and very large LCDs are appearing as "dynamic" wallpaper in museum display settings.

EXERCISE

3.1 Calculate the intrinsic carrier density n_i for silicon at $T = 50$ K and 350 K.
Ans. $9.6 \times 10^{-39}/\text{cm}^3$; $4.15 \times 10^{11}/\text{cm}^3$

3.2 Doped Semiconductors

The intrinsic silicon crystal described above has equal concentrations of free electrons and holes, generated by thermal generation. These concentrations are far too small for silicon to conduct appreciable current at room temperature. Also, the carrier concentrations and hence the conductivity are strong functions of temperature, not a desirable property in an electronic device. Fortunately, a method was developed to change the carrier concentration in a semiconductor crystal substantially and in a precisely controlled manner. This process is known as doping, and the resulting silicon is referred to as **doped silicon**.

Doping involves introducing impurity atoms into the silicon crystal in sufficient numbers to substantially increase the concentration of either free electrons or holes but with little or no change in the crystal properties of silicon. To increase the concentration of free electrons, n, silicon is doped with an element with a valence of 5, such as phosphorus. The resulting doped silicon is then said to be of ***n* type**. To increase the concentration of holes, p, silicon is doped with an element having a valence of 3, such as boron, and the resulting doped silicon is said to be of ***p* type**.

Figure 3.3 shows a silicon crystal doped with phosphorus impurity. The dopant (phosphorus) atoms replace some of the silicon atoms in the crystal structure. Since the phosphorus atom has five electrons in its outer shell, four of these electrons form covalent bonds with the neighboring atoms, and the fifth electron becomes a free electron. Thus each phosphorus atom *donates* a free electron to the silicon crystal, and the phosphorus impurity is called a **donor**. It should be clear, though, that no holes are generated by this process. The net positive charge associated with the phosphorus atom is a **bound charge** that does not move through the crystal.

If the concentration of donor atoms is N_D, where N_D is usually much greater than n_i, the concentration of free electrons in the n-type silicon will be

$$n_n \simeq N_D \tag{3.4}$$

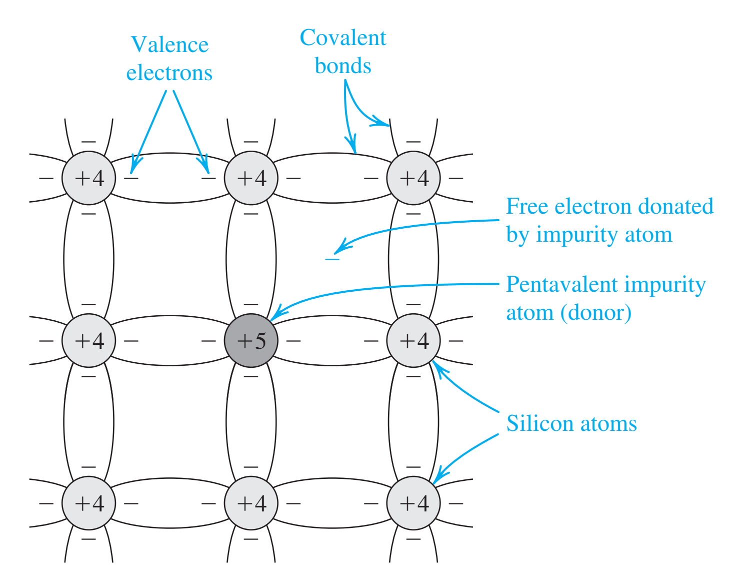


Figure 3.3 A silicon crystal doped by a pentavalent element. Each dopant atom donates a free electron and is thus called a donor. The doped semiconductor becomes *n* type.

where the subscript *n* denotes *n*-type silicon. Thus n_n is determined by the doping concentration and not by temperature. This is not the case, however, for the hole concentration. All the holes in the *n*-type silicon are those generated by thermal ionization. Their concentration p_n can be found by noting that the relationship in Eq. (3.3) applies equally well for doped silicon, provided thermal equilibrium is achieved. Thus for *n*-type silicon

$$p_n n_n = n_i^2$$

Substituting for n_n from Eq. (3.4), we obtain for p_n

$$p_n \simeq \frac{n_i^2}{N_D} \tag{3.5}$$

Thus p_n will have the same dependence on temperature as that of n_i^2. Finally, we note that in *n*-type silicon the concentration of free electrons n_n will be much larger than that of holes. Hence electrons are said to be the **majority** charge carriers and holes the **minority** charge carriers in *n*-type silicon.

To obtain *p*-type silicon in which holes are the majority charge carriers, a trivalent impurity such as boron is used. Figure 3.4 shows a silicon crystal doped with boron. Note that the boron atoms replace some of the silicon atoms in the silicon crystal structure. Since each boron atom has three electrons in its outer shell, it **accepts** an electron from a neighboring atom, thus forming covalent bonds. The result is a hole in the neighboring atom and a bound negative charge at the **acceptor** (boron) atom. It follows that each acceptor atom provides a hole. If the acceptor doping concentration is N_A, where $N_A \gg n_i$, the hole concentration becomes

$$p_p \simeq N_A \tag{3.6}$$

where the subscript *p* denotes *p*-type silicon. Thus, here the majority carriers are holes and their concentration is determined by N_A. The concentration of minority electrons can be found

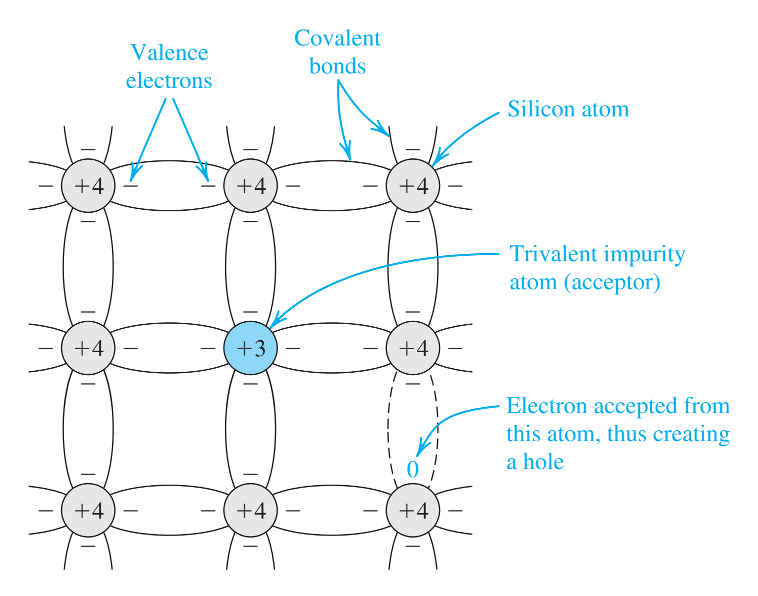


Figure 3.4 A silicon crystal doped with boron, a trivalent impurity. Each dopant atom gives rise to a hole, and the semiconductor becomes *p* type.

by using the relationship

$$p_p n_p = n_i^2$$

and substituting for p_p from Eq. (3.6),

$$n_p \simeq \frac{n_i^2}{N_A} \tag{3.7}$$

Thus, the concentration of the minority electrons will have the same temperature dependence as that of n_i^2.

It should be emphasized that a piece of *n*-type or *p*-type silicon is electrically neutral; the charge of the majority free carriers (electrons in the *n*-type and holes in the *p*-type silicon) are neutralized by the bound charges associated with the impurity atoms.

Example 3.2

Consider an *n*-type silicon for which the dopant concentration $N_D = 10^{17}/\text{cm}^3$. Find the electron and hole concentrations at $T = 300$ K.

Solution

The concentration of the majority electrons is

$$n_n \simeq N_D = 10^{17}/\text{cm}^3$$

Example 3.2 *continued*

The concentration of the minority holes is

$$p_n \simeq \frac{n_i^2}{N_D}$$

In Example 3.1 we found that at $T = 300$ K, $n_i = 1.5 \times 10^{10}/\text{cm}^3$. Thus,

$$p_n = \frac{\left(1.5 \times 10^{10}\right)^2}{10^{17}}$$
$$= 2.25 \times 10^3/\text{cm}^3$$

Observe that $n_n \gg n_i$ and that n_n is vastly higher than p_n.

EXERCISES

3.2 For the situation in Example 3.2, find the electron and hole concentrations at 350 K. You may use the value of n_i at $T = 350$ K found in Exercise 3.1.
Ans. $n_n = 10^{17}/\text{cm}^3$, $p_n = 1.72 \times 10^6/\text{cm}^3$

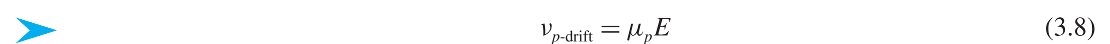
3.3 For a silicon crystal doped with boron, what must N_A be if at $T = 300$ K the electron concentration drops below the intrinsic level by a factor of 10^6?
Ans. $N_A = 1.5 \times 10^{16}/\text{cm}^3$

3.3 Current Flow in Semiconductors

There are two distinctly different mechanisms for the movement of charge carriers and hence for current flow in semiconductors: drift and diffusion.

3.3.1 Drift Current

When an electrical field E is established in a semiconductor crystal, holes are accelerated in the direction of E, and free electrons are accelerated in the direction opposite to that of E. This situation is illustrated in Fig. 3.5. The holes acquire a velocity $v_{p\text{-drift}}$ given by



$$v_{p\text{-drift}} = \mu_p E \tag{3.8}$$

where μ_p is a constant called the **hole mobility**: It represents the degree of ease by which holes move through the silicon crystal in response to the electrical field E. Since velocity has the units of centimeters per second and E has the units of volts per centimeter, we see from Eq. (3.8) that the mobility μ_p must have the units of centimeters squared per volt-second (cm^2/V · s). For intrinsic silicon $\mu_p = 480$ cm^2/V · s.

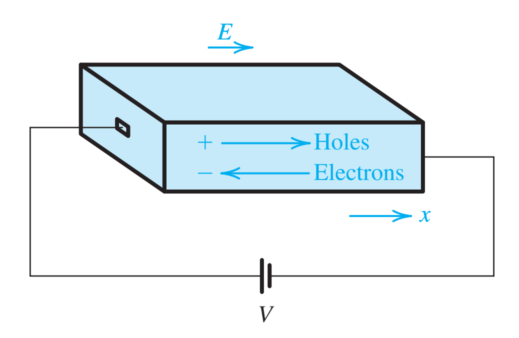


Figure 3.5 An electric field E established in a bar of silicon causes the holes to drift in the direction of E and the free electrons to drift in the opposite direction. Both the hole and electron drift currents are in the direction of E.

The free electrons acquire a drift velocity $v_{n\text{-drift}}$ given by

$$v_{n\text{-drift}} = -\mu_n E \tag{3.9}$$

where the result is negative because the electrons move in the direction opposite to E. Here μ_n is the **electron mobility**, which for intrinsic silicon is about 1350 cm^2/V · s. Note that μ_n is about 2.5 times μ_p, signifying that electrons move with much greater ease through the silicon crystal than do holes.

Let's now return to the single-crystal silicon bar shown in Fig. 3.5. Let the concentration of holes be p and that of free electrons n. We wish to calculate the current component due to the flow of holes. Consider a plane perpendicular to the x direction. In one second, the hole charge that crosses that plane will be $(Aqpv_{p\text{-drift}})$ coulombs, where A is the cross-sectional area of the silicon bar and q is the magnitude of electron charge. This then must be the hole component of the drift current flowing through the bar,

$$I_p = Aqpv_{p\text{-drift}} \tag{3.10}$$

Substituting for $v_{p\text{-drift}}$ from Eq. (3.8), we obtain

$$I_p = Aqp\mu_p E$$

We are usually interested in the current density J_p, which is the current per unit cross-sectional area,

$$J_p = \frac{I_p}{A} = qp\mu_p E \tag{3.11}$$

The current component due to the drift of free electrons can be found in a similar manner. Note, however, that electrons drifting from right to left result in a current component from left to right. This is because of the convention of taking the direction of current flow as the direction of flow of positive charge and opposite to the direction of flow of negative charge. Thus,

$$I_n = -Aqnv_{n\text{-drift}}$$

Substituting for $v_{n\text{-drift}}$ from Eq. (3.9), we obtain the current density $J_n = I_n/A$ as

$$J_n = qn\mu_n E \tag{3.12}$$

The total drift current density can now be found by summing J_p and J_n from Eqs. (3.11) and (3.12),

$$J = J_p + J_n = q(p\mu_p + n\mu_n)E \tag{3.13}$$

This relationship can be written as

$$J = \sigma E \tag{3.14}$$

or

$$J = E/\rho \tag{3.15}$$

where the **conductivity** σ is given by

$$\sigma = q(p\mu_p + n\mu_n) \tag{3.16}$$

and the **resistivity** ρ is given by



$$\rho \equiv \frac{1}{\sigma} = \frac{1}{q(p\mu_p + n\mu_n)} \tag{3.17}$$

Observe that Eq. (3.15) is a form of Ohm's law and can be written alternately as

$$\rho = \frac{E}{J} \tag{3.18}$$

Thus the units of ρ are obtained from: $\frac{\text{V/cm}}{\text{A/cm}^2} = \Omega \cdot \text{cm}$.

Example 3.3

Find the resistivity of (a) intrinsic silicon and (b) p-type silicon with $N_A = 10^{16}/\text{cm}^3$. Use $n_i = 1.5 \times 10^{10}/\text{cm}^3$, and assume that for intrinsic silicon $\mu_n = 1350 \text{ cm}^2/\text{V} \cdot \text{s}$ and $\mu_p = 480 \text{ cm}^2/\text{V} \cdot \text{s}$, and for the doped silicon $\mu_n = 1110 \text{ cm}^2/\text{V} \cdot \text{s}$ and $\mu_p = 400 \text{ cm}^2/\text{V} \cdot \text{s}$. (Note that doping results in reduced carrier mobilities.)

Solution

(a) For intrinsic silicon,

$$p = n = n_i = 1.5 \times 10^{10}/\text{cm}^3$$

Thus,

$$\rho = \frac{1}{q(p\mu_p + n\mu_n)}$$

$$\rho = \frac{1}{1.6 \times 10^{-19}(1.5 \times 10^{10} \times 480 + 1.5 \times 10^{10} \times 1350)}$$

$$= 2.28 \times 10^5 \ \Omega \cdot \text{cm}$$

(b) For the p-type silicon

$$p_p \simeq N_A = 10^{16}/\text{cm}^3$$

$$n_p \simeq \frac{n_i^2}{N_A} = \frac{(1.5 \times 10^{10})^2}{10^{16}} = 2.25 \times 10^4/\text{cm}^3$$

Thus,

$$\rho = \frac{1}{q(p\mu_p + n\mu_n)}$$

$$= \frac{1}{1.6 \times 10^{-19}(10^{16} \times 400 + 2.25 \times 10^4 \times 1110)}$$

$$\simeq \frac{1}{1.6 \times 10^{-19} \times 10^{16} \times 400} = 1.56\ \Omega \cdot \text{cm}$$

Observe that the resistivity of the p-type silicon is determined almost entirely by the doping concentration. Also observe that doping the silicon reduces its resistivity by a factor of about 10^4, a truly remarkable change.

EXERCISE

3.4 A uniform bar of n-type silicon of 2-μm length has a voltage of 1 V applied across it. If $N_D = 10^{16}/\text{cm}^3$ and $\mu_n = 1350\ \text{cm}^2/\text{V} \cdot \text{s}$, find (a) the electron drift velocity, (b) the time it takes an electron to cross the 2-μm length, (c) the drift-current density, and (d) the drift current in the case that the silicon bar has a cross-sectional area of 0.25 μm^2.
Ans. 6.75×10^6 cm/s; 30 ps; 1.08×10^4 A/cm^2; 27 μA

3.3.2 Diffusion Current

Carrier diffusion occurs when the density of charge carriers in a piece of semiconductor is not uniform. For instance, if by some mechanism the concentration of, say, holes, is made higher in one part of a piece of silicon than in another, then holes will diffuse from the region of high concentration to the region of low concentration. Such a diffusion process is like that observed if one drops a few ink drops in a water-filled tank. The diffusion of charge carriers gives rise to a net flow of charge, or **diffusion current**.

As an example, consider the bar of silicon shown in Fig. 3.6(a): By some unspecified process, we have arranged to inject holes into its left side. This continuous hole injection gives rise to and maintains a hole **concentration profile** such as that shown in Fig. 3.6(b). This profile in turn causes holes to diffuse from left to right along the silicon bar, resulting in a hole current in the x direction. The magnitude of the current at any point is proportional to the slope of the concentration profile, or the **concentration gradient**, at that point,

$$J_p = -qD_p \frac{dp(x)}{dx} \tag{3.19}$$

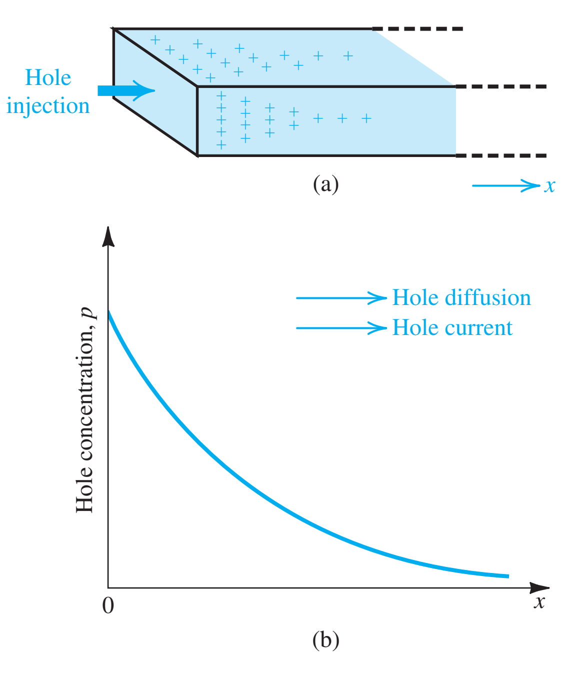


Figure 3.6 A bar of silicon **(a)** into which holes are injected, thus creating the hole concentration profile along the x axis, shown in **(b)**. The holes diffuse in the positive direction of x and give rise to a hole diffusion current in the same direction. Note that we are not showing the circuit to which the silicon bar is connected.

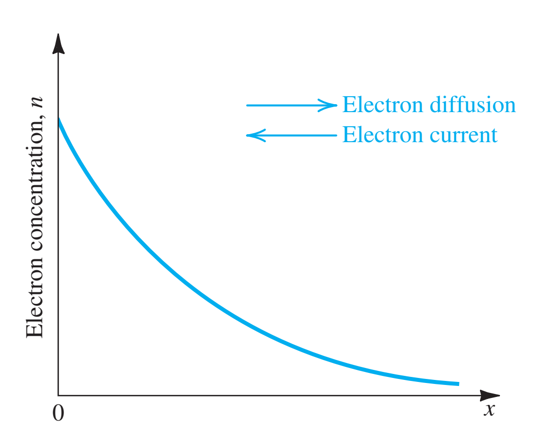


Figure 3.7 If the electron concentration profile shown is established in a bar of silicon, electrons diffuse in the x direction, giving rise to an electron diffusion current in the negative-x direction.

where J_p is the hole-current density (A/cm^2), q is the magnitude of electron charge, D_p is a constant called the **diffusion constant** or **diffusivity** of holes; and $p(x)$ is the hole concentration at point x. Note that the gradient (dp/dx) is negative, resulting in a positive current in the x direction, as should be expected.

In the case of electron diffusion resulting from an electron concentration gradient (see Fig. 3.7), a similar relationship applies, giving the electron-current density,



$$J_n = qD_n \frac{dn(x)}{dx} \tag{3.20}$$

where D_n is the diffusion constant or diffusivity of electrons. Observe that a negative (dn/dx) gives rise to a negative current, a result of the convention that the positive direction of current is taken to be that of the flow of positive charge (and opposite to that of the flow of negative

charge). For holes and electrons diffusing in intrinsic silicon, typical values for the diffusion constants are $D_p = 12\ \text{cm}^2/\text{s}$ and $D_n = 35\ \text{cm}^2/\text{s}$.

At this point the reader is probably wondering where the diffusion current in the silicon bar in Fig. 3.6(a) goes. A good question, as we are not showing how the right-side end of the bar is connected to the rest of the circuit. We will address this and related questions in detail in our discussion of the *pn* junction in later sections.

Example 3.4

Consider a bar of silicon in which a hole concentration profile described by

$$p(x) = p_0 e^{-x/L_p}$$

is established. Find the hole-current density at $x = 0$. Let $p_0 = 10^{16}/\text{cm}^3$, $L_p = 1\ \mu\text{m}$, and $D_p = 12\ \text{cm}^2/\text{s}$. If the cross-sectional area of the bar is $100\ \mu\text{m}^2$, find the current I_p.

Solution

$$\begin{aligned} J_p &= -qD_p \frac{dp(x)}{dx} \\ &= -qD_p \frac{d}{dx}\left[p_0 e^{-x/L_p}\right] \\ &= q\frac{D_p}{L_p} p_0 e^{-x/L_p} \end{aligned}$$

Thus,

$$\begin{aligned} J_p(0) &= q\frac{D_p}{L_p} p_0 \\ &= 1.6 \times 10^{-19} \times \frac{12}{1 \times 10^{-4}} \times 10^{16} \\ &= 192\ \text{A/cm}^2 \end{aligned}$$

The current I_p can be found from

$$\begin{aligned} I_p &= J_p \times A \\ &= 192 \times 100 \times 10^{-8} \\ &= 192\ \mu\text{A} \end{aligned}$$

EXERCISE

3.5 The linear electron-concentration profile shown in Fig. E3.5 has been established in a piece of silicon. If $n_0 = 10^{17}/\text{cm}^3$ and $W = 1\ \mu\text{m}$, find the electron-current density in microamperes per micron squared ($\mu\text{A}/\mu\text{m}^2$). If a diffusion current of 1 mA is required, what must the cross-sectional area (in a direction perpendicular to the page) be? Recall that $D_n = 35\ \text{cm}^2/\text{s}$.

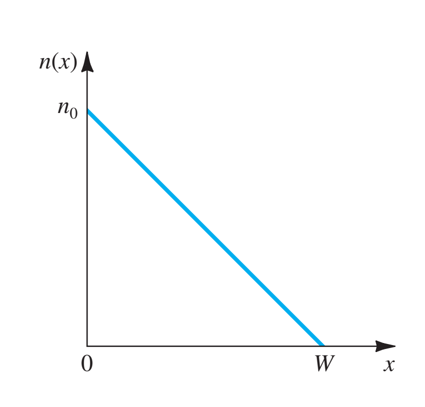


Figure E3.5

Ans. 56 μA/μm^2; 18 μm^2

3.3.3 Relationship between D and μ

A simple but powerful relationship ties the diffusion constant with the mobility,



$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = V_T \tag{3.21}$$

where $V_T = kT/q$. The parameter V_T is known as the **thermal voltage**. At room temperature, $T \simeq 300$ K and $V_T = 25.9$ mV. We will encounter V_T repeatedly throughout this book. The relationship in Eq. (3.21) is known as the **Einstein relationship**.

EXERCISE

3.6 Use the Einstein relationship to find D_n and D_p for intrinsic silicon using $\mu_n = 1350$ cm^2/V·s and $\mu_p = 480$ cm^2/V·s.
Ans. 35 cm^2/s; 12.4 cm^2/s

3.4 The *pn* Junction

Having learned important semiconductor concepts, we are now ready to consider our first practical semiconductor structure—the *pn* junction. As mentioned previously, the *pn* junction implements the diode (Chapter 4) and plays the dominant role in the structure and operation of the bipolar junction transistor (BJT, Chapter 6). As well, understanding *pn* junctions is very important to the study of the MOSFET operation (Chapter 5).

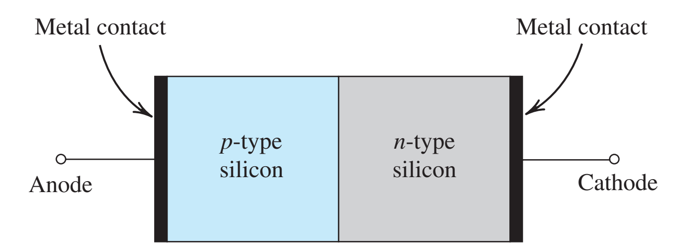


Figure 3.8 Simplified physical structure of the *pn* junction. (Actual geometries are given in Appendix A.) As the *pn* junction implements the junction diode, its terminals are labeled anode and cathode.

3.4.1 Physical Structure

Figure 3.8 shows a simplified physical structure of the *pn* junction. It consists of a *p*-type semiconductor (e.g., silicon) brought into close contact with an *n*-type semiconductor material (also silicon). In actual practice, both the *p* and *n* regions are part of the same silicon crystal; that is, the *pn* junction is formed within a single silicon crystal by creating regions of different dopings (*p* and *n* regions). Appendix A provides a description of the fabrication process of integrated circuits including *pn* junctions. As indicated in Fig. 3.8, external wire connections are made to the *p* and *n* regions through metal (aluminum) contacts. If the *pn* junction is used as a diode, these constitute the diode terminals and are therefore labeled "anode" and "cathode" in keeping with diode terminology.[3]

3.4.2 Operation with Open-Circuit Terminals

Figure 3.9 shows a *pn* junction under open-circuit conditions—that is, the external terminals are left open. The "+" signs in the *p*-type material denote the majority holes. The charge of these holes is neutralized by an equal amount of bound negative charge associated with the acceptor atoms. For simplicity, these bound charges are not shown in the diagram. Also not shown are the minority electrons generated in the *p*-type material by thermal ionization.

In the *n*-type material the majority electrons are indicated by "–" signs. Here also, the bound positive charge, which neutralizes the charge of the majority electrons, is not shown in order to keep the diagram simple. The *n*-type material also contains minority holes generated by thermal ionization but not shown in the diagram.

The Diffusion Current I_D Because the concentration of holes is high in the *p* region and low in the *n* region, holes diffuse across the junction from the *p* side to the *n* side. Similarly, electrons diffuse across the junction from the *n* side to the *p* side. These two current components add together to form the diffusion current I_D, whose direction is from the *p* side to the *n* side, as indicated in Fig. 3.9.

The Depletion Region The holes that diffuse across the junction into the *n* region quickly recombine with some of the majority electrons present there and thus disappear from the scene. This recombination process results also in the disappearance of some free electrons from the

[3]This terminology in fact is a carryover from that used with vacuum-tube technology, which was the technology for making diodes and other electronic devices until the invention of the transistor in 1947. This event ushered in the era of solid-state electronics, which changed not only electronics, communications, and computers but indeed the world!

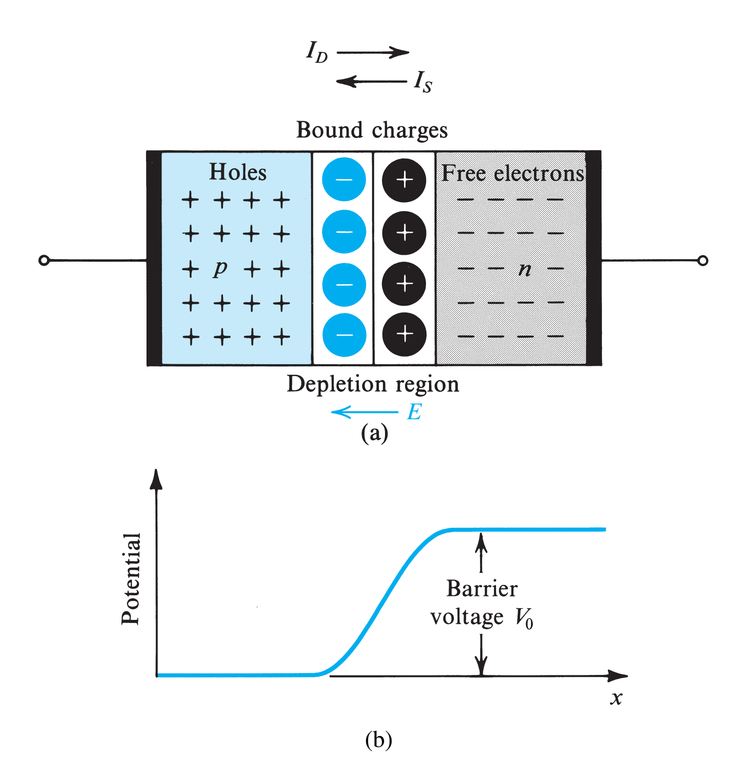


Figure 3.9 **(a)** The *pn* junction with no applied voltage (open-circuited terminals). **(b)** The potential distribution along an axis perpendicular to the junction.

n-type material. Thus some of the bound positive charge will no longer be neutralized by free electrons, and this charge is said to have been **uncovered**. Since recombination takes place close to the junction, there will be a region close to the junction that is *depleted of free electrons* and contains uncovered bound positive charge, as indicated in Fig. 3.9.

The electrons that diffuse across the junction into the *p* region quickly recombine with some of the majority holes there, and thus disappear from the scene. This results also in the disappearance of some majority holes, causing some of the bound negative charge to be uncovered (i.e., no longer neutralized by holes). Thus, in the *p* material close to the junction, there will be a region *depleted of holes* and containing uncovered bound negative charge, as indicated in Fig. 3.9.

From the above it follows that a **carrier-depletion region** will exist on both sides of the junction, with the *n* side of this region positively charged and the *p* side negatively charged. This carrier-depletion region—or, simply, **depletion region**—is also called the **space-charge region**. The charges on both sides of the depletion region cause an electric field E to be established across the region in the direction indicated in Fig. 3.9. Hence a potential difference results across the depletion region, with the *n* side at a positive voltage relative to the *p* side, as shown in Fig. 3.9(b). Thus the resulting electric field opposes the diffusion of holes into the *n* region and electrons into the *p* region. In fact, the voltage drop across the depletion region acts as a **barrier** that has to be overcome for holes to diffuse into the *n* region and electrons to diffuse into the *p* region. The larger the barrier voltage, the smaller the number of carriers that will be able to overcome the barrier, and hence the lower the magnitude of diffusion current. Thus it is the appearance of the barrier voltage V_0 that limits the carrier diffusion process. It follows that the diffusion current I_D depends strongly on the voltage drop V_0 across the depletion region.

The Drift Current I_S and Equilibrium In addition to the current component I_D due to majority-carrier diffusion, a component due to minority-carrier drift exists across the junction. Specifically, some of the thermally generated holes in the *n* material move toward the junction and reach the edge of the depletion region. There, they experience the electric field in the depletion region, which sweeps them across that region into the *p* side. Similarly, some of the minority thermally generated electrons in the *p* material move to the edge of the depletion region and get swept by the electric field in the depletion region across that region into the *n* side. These two current components—electrons moved by drift from *p* to *n* and holes moved by drift from *n* to *p*—add together to form the drift current I_S, whose direction is from the *n* side to the *p* side of the junction, as indicated in Fig. 3.9. Since the current I_S is carried by thermally generated minority carriers, its value is strongly dependent on temperature; however, it is independent of the value of the depletion-layer voltage V_0. This is due to the fact that the drift current is determined by the number of minority carriers that make it to the edge of the depletion region; any minority carriers that manage to get to the edge of the depletion region will be swept across by E irrespective of the value of E or, correspondingly, of V_0.

Under open-circuit conditions (Fig. 3.9) no external current exists; thus the two opposite currents across the junction must be equal in magnitude:

$$I_D = I_S$$

This equilibrium condition[4] is maintained by the barrier voltage V_0. Thus, if for some reason I_D exceeds I_S, then more bound charge will be uncovered on both sides of the junction, the depletion layer will widen, and the voltage across it (V_0) will increase. This in turn causes I_D to decrease until equilibrium is achieved with $I_D = I_S$. On the other hand, if I_S exceeds I_D, then the amount of uncovered charge will decrease, the depletion layer will narrow, and the voltage across it (V_0) will decrease. This causes I_D to increase until equilibrium is achieved with $I_D = I_S$.

The Junction Built-in Voltage With no external voltage applied, the barrier voltage V_0 across the *pn* junction can be shown to be given by[5]

$$V_0 = V_T \ln\left(\frac{N_A N_D}{n_i^2}\right) \tag{3.22}$$

where N_A and N_D are the doping concentrations of the *p* side and *n* side of the junction, respectively. Thus V_0 depends both on doping concentrations and on temperature. It is known as the **junction built-in voltage**. Typically, for silicon at room temperature, V_0 is in the range of 0.6 V to 0.9 V.

When the *pn* junction terminals are left open-circuited, the voltage measured between them will be zero. That is, the voltage V_0 across the depletion region *does not* appear between the junction terminals. This is because of the contact voltages existing at the metal–semiconductor junctions at the terminals, which counter and exactly balance the barrier voltage. If this were not the case, we would have been able to draw energy from the isolated *pn* junction, which would clearly violate the principle of conservation of energy.

Width of and Charge Stored in the Depletion Region Figure 3.10 provides further illustration of the situation that obtains in the *pn* junction when the junction is in equilibrium.

---

[4]In fact, in equilibrum the equality of drift and diffusion currents applies not just to the total currents but also to their individual components. That is, the hole drift current must equal the hole diffusion current and, similarly, the electron drift current must equal the electron diffusion current.

[5]The derivation of this formula and of a number of others in this chapter can be found in textbooks dealing with devices, such as that by Streetman and Bannerjee (see the reading list in Appendix I).

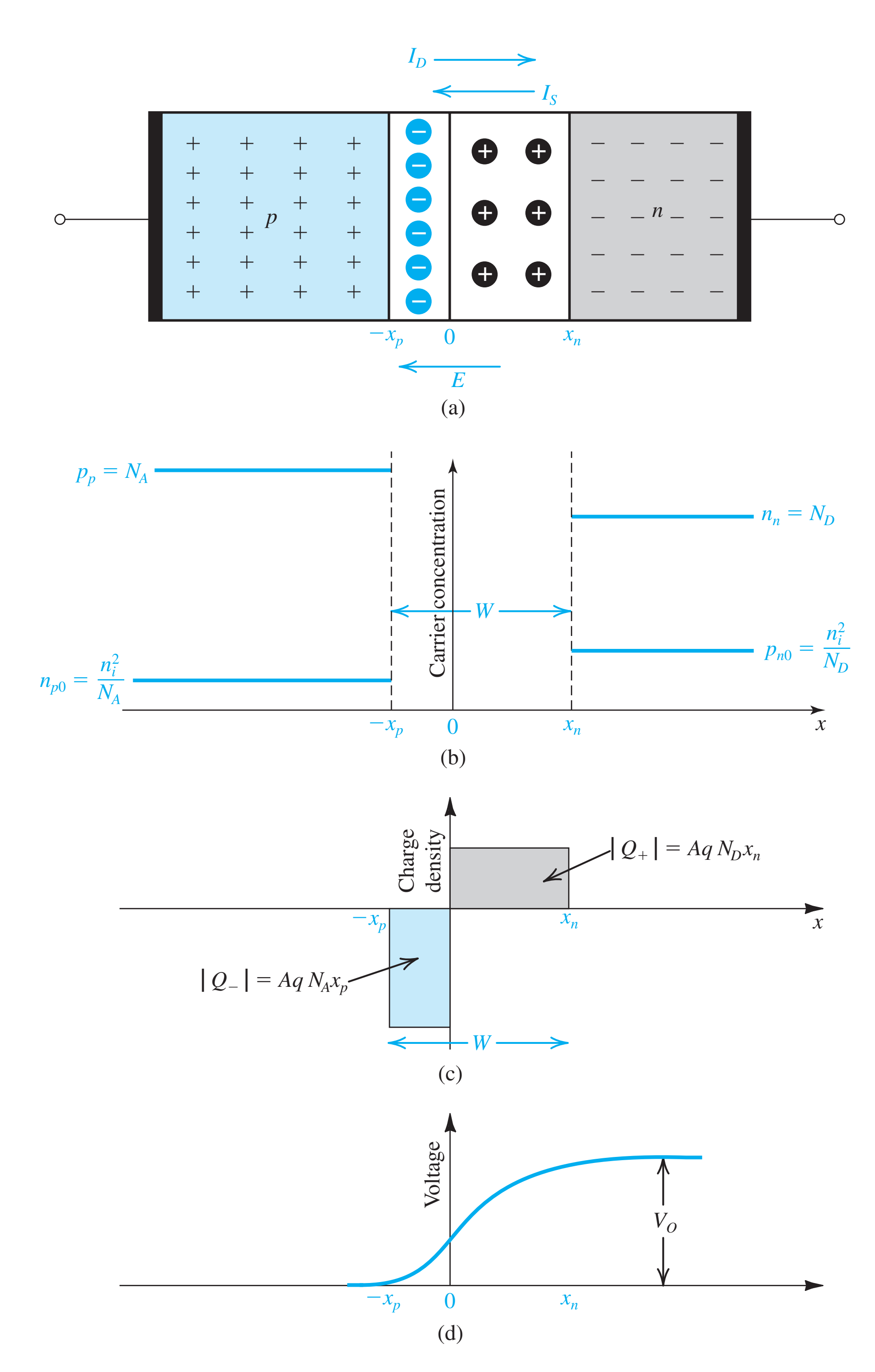


Figure 3.10 **(a)** A *pn* junction with the terminals open-circuited. **(b)** Carrier concentrations; note that $N_A > N_D$. **(c)** The charge stored in both sides of the depletion region; $Q_J = |Q_+| = |Q_-|$. **(d)** The built-in voltage V_0.

In Fig. 3.10(a) we show a junction in which $N_A > N_D$, a typical situation in practice. This is borne out by the carrier concentration on both sides of the junction, as shown in Fig. 3.10(b). Note that we have denoted the minority-carrier concentrations in both sides by n_{p0} and p_{n0}, with the additional subscript "0" signifying equilibrium (i.e., before external voltages are applied, as will be seen in the next section). Observe that the depletion region extends in both the *p* and *n* materials and that equal amounts of charge exist on both sides (Q_+ and Q_- in Fig. 3.10c). However, since usually unequal dopings N_A and N_D are used, as in the case illustrated in Fig. 3.10, the width of the depletion layer will not be the same on the two sides. Rather, to uncover the same amount of charge, the depletion layer will extend deeper into the more lightly doped material. Specifically, if we denote the width of the depletion region in the *p* side by x_p and in the *n* side by x_n, we can express the magnitude of the charge on the *n* side of the junction as

$$|Q_+| = qAx_nN_D \tag{3.23}$$

and that on the *p* side of the junction as

$$|Q_-| = qAx_pN_A \tag{3.24}$$

where *A* is the cross-sectional area of the junction in the plane perpendicular to the page. The charge equality condition can now be written as

$$qAx_nN_D = qAx_pN_A$$

which can be rearranged to yield

$$\frac{x_n}{x_p} = \frac{N_A}{N_D} \tag{3.25}$$

In actual practice, it is usual for one side of the junction to be much more heavily doped than the other, with the result that the depletion region exists almost entirely on one side (the lightly doped side).

The width *W* of the depletion layer can be shown to be given by

$$W = x_n + x_p = \sqrt{\frac{2\epsilon_s}{q}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)V_0} \tag{3.26}$$

where ϵ_s is the electrical permittivity of silicon $= 11.7\epsilon_0 = 11.7 \times 8.85 \times 10^{-14}$ F/cm $= 1.04 \times 10^{-12}$ F/cm. Typically *W* is in the range 0.1 μm to 1 μm. Eqs. (3.25) and (3.26) can be used to obtain x_n and x_p in terms of *W* as

$$x_n = W\frac{N_A}{N_A + N_D} \tag{3.27}$$

$$x_p = W\frac{N_D}{N_A + N_D} \tag{3.28}$$

The charge stored on either side of the depletion region can be expressed in terms of *W* by utilizing Eqs. (3.23) and (3.27) to obtain

$$Q_J = |Q_+| = |Q_-|$$

$$Q_J = Aq\left(\frac{N_AN_D}{N_A + N_D}\right)W \tag{3.29}$$

Finally, we can substitute for *W* from Eq. (3.26) to obtain

$$Q_J = A\sqrt{2\epsilon_s q\left(\frac{N_AN_D}{N_A + N_D}\right)V_0} \tag{3.30}$$

These expressions for Q_J will prove useful in subsequent sections.

Example 3.5

Consider a *pn* junction in equilibrium at room temperature ($T = 300$ K) for which the doping concentrations are $N_A = 10^{18}/\text{cm}^3$ and $N_D = 10^{16}/\text{cm}^3$ and the cross-sectional area $A = 10^{-4}$ cm^2. Calculate p_p, n_{p0}, n_n, p_{n0}, V_0, W, x_n, x_p, and Q_J. Use $n_i = 1.5 \times 10^{10}/\text{cm}^3$.

Solution

$$p_p \simeq N_A = 10^{18} \text{ cm}^{-3}$$

$$n_{p0} = \frac{n_i^2}{p_p} \simeq \frac{n_i^2}{N_A} = \frac{(1.5 \times 10^{10})^2}{10^{18}} = 2.25 \times 10^2 \text{ cm}^{-3}$$

$$n_n \simeq N_D = 10^{16} \text{ cm}^{-3}$$

$$p_{n0} = \frac{n_i^2}{n_n} \simeq \frac{n_i^2}{N_D} = \frac{(1.5 \times 10^{10})^2}{10^{16}} = 2.25 \times 10^4 \text{ cm}^{-3}$$

To find V_0 we use Eq. (3.22),

$$V_O = V_T \ln\left(\frac{N_A N_D}{n_i^2}\right)$$

where

$$V_T = \frac{kT}{q} = \frac{8.62 \times 10^{-5} \times 300 \text{ (eV)}}{q \qquad \text{(e)}}$$
$$= 25.9 \times 10^{-3} \text{ V}$$

Thus,

$$V_0 = 25.9 \times 10^{-3} \ln\left(\frac{10^{18} \times 10^{16}}{2.25 \times 10^{20}}\right)$$
$$= 0.814 \text{ V}$$

To determine W we use Eq. (3.26):

$$W = \sqrt{\frac{2 \times 1.04 \times 10^{-12}}{1.6 \times 10^{-19}}\left(\frac{1}{10^{18}} + \frac{1}{10^{16}}\right) \times 0.814}$$
$$= 3.27 \times 10^{-5} \text{ cm} = 0.327 \text{ µm}$$

To determine x_n and x_p we use Eqs. (3.27) and (3.28), respectively:

$$x_n = W\frac{N_A}{N_A + N_D}$$
$$= 0.327\frac{10^{18}}{10^{18} + 10^{16}} = 0.324 \text{ µm}$$
$$x_p = W\frac{N_D}{N_A + N_D}$$
$$= 0.327\frac{10^{16}}{10^{18} + 10^{16}} = 0.003 \text{ µm}$$

Finally, to determine the charge stored on either side of the depletion region, we use Eq. (3.29):

$$Q_J = 10^{-4} \times 1.6 \times 10^{-19}\left(\frac{10^{18} \times 10^{16}}{10^{18} + 10^{16}}\right) \times 0.327 \times 10^{-4}$$
$$= 5.18 \times 10^{-12} \text{ C} = 5.18 \text{ pC}$$

EXERCISES

3.7 Show that

$$V_0 = \frac{1}{2}\left(\frac{q}{\epsilon_s}\right)\left(\frac{N_A N_D}{N_A + N_D}\right)W^2$$

3.8 Show that for a *pn* junction in which the *p* side is much more heavily doped than the *n* side (i.e., $N_A \gg N_D$), referred to as a p^+n diode, Eqs. (3.26), (3.27), (3.28), (3.29), and (3.30) can be simplified as follows:

$$W \simeq \sqrt{\frac{2\epsilon_s}{qN_D}V_0} \qquad (3.26')$$

$$x_n \simeq W \qquad (3.27')$$

$$x_p \simeq W/(N_A/N_D) \qquad (3.28')$$

$$Q_J \simeq AqN_D W \qquad (3.29')$$

$$Q_J \simeq A\sqrt{2\epsilon_s qN_D V_0} \qquad (3.30')$$

3.9 If in the fabrication of the *pn* junction in Example 3.5, it is required to increase the minority-carrier concentration in the *n* region by a factor of 2, what must be done?
Ans. Lower N_D by a factor of 2.

3.5 The *pn* Junction with an Applied Voltage

Having studied the open-circuited *pn* junction in detail, we are now ready to apply a dc voltage between its two terminals to find its electrical conduction properties. If the voltage is applied so that the *p* side is made more positive than the *n* side, it is referred to as a forward-bias[6] voltage. Conversely, if our applied dc voltage is such that it makes the *n* side more positive than the *p* side, it is said to be a reverse-bias voltage. As will be seen, the *pn* junction exhibits vastly different conduction properties in its forward and reverse directions.

Our plan is as follows. We begin by a simple qualitative description in Section 3.5.1 and then consider an analytical description of the *i*–*v* characteristic of the junction in Section 3.5.2.

3.5.1 Qualitative Description of Junction Operation

Figure 3.11 shows the *pn* junction under three different conditions: (a) the open-circuit or equilibrium condition studied in the previous section; (b) the reverse-bias condition, where a dc voltage V_R is applied; and (c) the forward-bias condition, where a dc voltage V_F is applied.

[6] For the time being, we take the term *bias* to refer simply to the application of a dc voltage. We will see in later chapters that it has a deeper meaning in the design of electronic circuits.

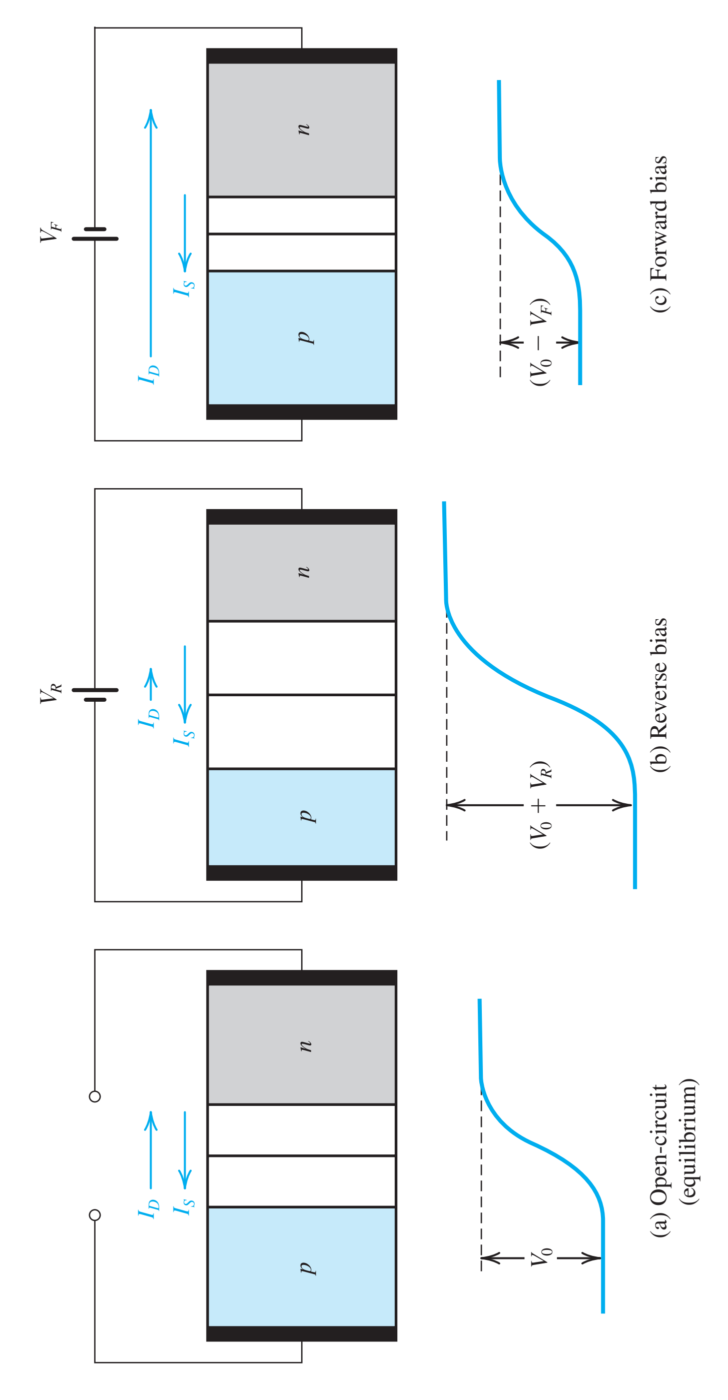


Figure 3.11 The *pn* junction in: **(a)** equilibrium; **(b)** reverse bias; **(c)** forward bias.

Observe that in the open-circuit case, a barrier voltage V_0 develops, making n more positive than p, and limiting the diffusion current I_D to a value exactly equal to the drift current I_S, thus resulting in a zero current at the junction terminals, as should be the case, since the terminals are open-circuited. Also, as mentioned previously, the barrier voltage V_0, though it establishes the current equilibrium across the junction, does *not* in fact appear between the junction terminals.

Consider now the reverse-bias case in (b). The externally applied reverse-bias voltage V_R is in the direction to add to the barrier voltage, and it does, thus increasing the effective barrier voltage to $(V_0 + V_R)$ as shown. This reduces the number of holes that diffuse into the n region and the number of electrons that diffuse into the p region. The end result is that the diffusion current I_D is dramatically reduced. As will be seen shortly, a reverse-bias voltage of a volt or so is sufficient to cause $I_D \simeq 0$, and the current across the junction and through the external circuit will be equal to I_S. Recalling that I_S is the current due to the drift across the depletion region of the thermally generated minority carriers, we expect I_S to be very small and to be strongly dependent on temperature. We will show this to be the case very shortly. We thus conclude that in the reverse direction, the *pn* junction conducts a very small and almost-constant current equal to I_S.

Before leaving the reverse-bias case, observe that the increase in barrier voltage will be accompanied by a corresponding increase in the stored uncovered charge on both sides of the depletion region. This in turn means a wider depletion region, needed to uncover the additional charge required to support the larger barrier voltage $(V_0 + V_R)$. Analytically, these results can be obtained easily by a simple extension of the results of the equilibrium case. Thus the width of the depletion region can be obtained by replacing V_0 in Eq. (3.26) by $(V_0 + V_R)$,

$$W = x_n + x_p = \sqrt{\frac{2\epsilon_s}{q}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)(V_0 + V_R)} \qquad (3.31)$$

and the magnitude of the charge stored on either side of the depletion region can be determined by replacing V_0 in Eq. (3.30) by $(V_0 + V_R)$,

$$Q_J = A\sqrt{2\epsilon_s q\left(\frac{N_A N_D}{N_A + N_D}\right)(V_0 + V_R)} \qquad (3.32)$$

We next consider the forward-bias case shown in Fig. 3.11(c). Here the applied voltage V_F is in the direction that subtracts from the built-in voltage V_0, resulting in a reduced barrier voltage $(V_0 - V_F)$ across the depletion region. This reduced barrier voltage will be accompanied by reduced depletion-region charge and correspondingly narrower depletion-region width W. Most importantly, the lowering of the barrier voltage will enable more holes to diffuse from p to n and more electrons to diffuse from n to p. Thus the diffusion current I_D increases substantially and, as will be seen shortly, can become many orders of magnitude larger than the drift current I_S. The current I in the external circuit is of course the difference between I_D and I_S,

$$I = I_D - I_S$$

and it flows in the forward direction of the junction, from p to n. We thus conclude that the *pn* junction can conduct a substantial current in the forward-bias region and that current is mostly a diffusion current whose value is determined by the forward-bias voltage V_F.

3.5.2 The Current–Voltage Relationship of the Junction

We are now ready to find an analytical expression that describes the current–voltage relationship of the *pn* junction. In the following we consider a junction operating with a forward applied voltage V and derive an expression for the current I that flows in the forward direction (from *p* to *n*). However, our derivation is general and will be seen to yield the reverse current when the applied voltage V is made negative.

From the qualitative description above we know that a forward-bias voltage V subtracts from the built-in voltage V_0, thus resulting in a lower barrier voltage $(V_0 - V)$. The lowered barrier in turn makes it possible for a greater number of holes to overcome the barrier and diffuse into the *n* region. A similar statement can be made about electrons from the *n* region diffusing into the *p* region.

Let us now consider the holes injected into the *n* region. The concentration of holes in the *n* region at the edge of the depletion region will increase considerably. In fact, an important result from device physics shows that the steady-state concentration at the edge of the depletion region will be

$$p_n(x_n) = p_{n0}e^{V/V_T} \tag{3.33}$$

That is, the concentration of the minority holes increases from the equilibrium value of p_{n0} (see Fig. 3.10) to the much larger value determined by the value of V, given by Eq. (3.33).

We describe this situation as follows: The forward-bias voltage V results in an **excess concentration** of minority holes at $x = x_n$, given by

$$\begin{aligned}\text{Excess concentration} &= p_{n0}e^{V/V_T} - p_{n0} \\ &= p_{n0}\left(e^{V/V_T} - 1\right)\end{aligned} \tag{3.34}$$

The increase in minority-carrier concentration in Eqs. (3.33) and (3.34) occurs at the edge of the depletion region $(x = x_n)$. As the injected holes diffuse into the *n* material, some will recombine with the majority electrons and disappear. Thus, the excess hole concentration will decay exponentially with distance. As a result, the total hole concentration in the *n* material will be given by

$$p_n(x) = p_{n0} + (\text{Excess concentration})e^{-(x-x_n)/L_p}$$

Substituting for the "Excess concentration" from Eq. (3.34) gives

$$p_n(x) = p_{n0} + p_{n0}\left(e^{V/V_T} - 1\right)e^{-(x-x_n)/L_p} \tag{3.35}$$

The exponential decay is characterized by the constant L_p, which is called the **diffusion length** of holes in the *n* material. The smaller the value of L_p, the faster the injected holes will recombine with the majority electrons, resulting in a steeper decay of minority-carrier concentration.

Figure 3.12 shows the steady-state minority-carrier concentration profiles on both sides of a *pn* junction in which $N_A \gg N_D$. Let's stay a little longer with the diffusion of holes into the *n* region. Note that the shaded region under the exponential represents the excess minority carriers (holes). From our study of diffusion in Section 3.3, we know that the establishment of a carrier concentration profile such as that in Fig. 3.12 is essential to support a steady-state diffusion current. In fact, we can now find the value of the hole–diffusion current density by applying Eq. (3.19),

$$J_p(x) = -qD_p\frac{dp_n(x)}{dx}$$

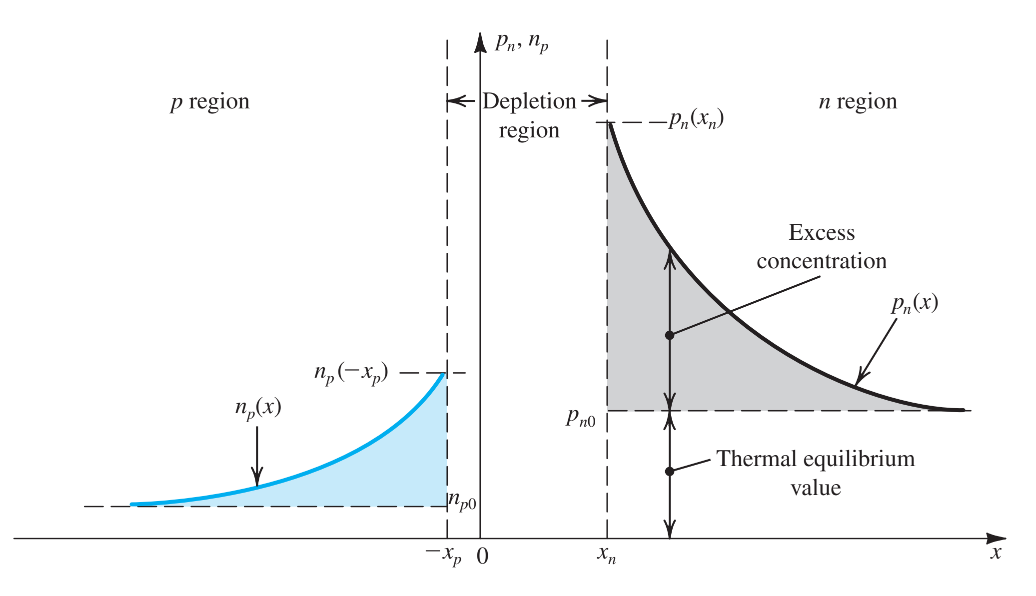


Figure 3.12 Minority-carrier distribution in a forward-biased *pn* junction. It is assumed that the *p* region is more heavily doped than the *n* region; $N_A \gg N_D$.

Substituting for $p_n(x)$ from Eq. (3.35) gives

$$J_p(x) = q\left(\frac{D_p}{L_p}\right)p_{n0}\left(e^{V/V_T} - 1\right)e^{-(x-x_n)/L_p} \tag{3.36}$$

As expected, $J_p(x)$ is highest at $x = x_n$,

$$J_p(x_n) = q\left(\frac{D_p}{L_p}\right)p_{n0}\left(e^{V/V_T} - 1\right) \tag{3.37}$$

and decays exponentially for $x > x_n$, as the minority holes recombine with the majority electrons. This recombination, however, means that the majority electrons will have to be replenished by a current that injects electrons from the external circuit into the *n* region of the junction. This latter current component has the same direction as the hole current (because electrons moving from right to left give rise to current in the direction from left to right). It follows that as $J_p(x)$ decreases, the electron current component increases by exactly the same amount, making the total current in the *n* material constant at the value given by Eq. (3.37).

An exactly parallel development can be applied to the electrons that are injected from the *n* to the *p* region, resulting in an electron diffusion current given by a simple adaptation of Eq. (3.37),

$$J_n(-x_p) = q\left(\frac{D_n}{L_n}\right)n_{p0}\left(e^{V/V_T} - 1\right) \tag{3.38}$$

Now, although the currents in Eqs. (3.37) and (3.38) are found at the two edges of the depletion region, their values do not change in the depletion region. Thus we can drop the location descriptors (x_n), $(-x_p)$, add the two current densities, and multiply by the junction area A to

obtain the total current I as

$$I = A(J_p + J_n)$$

$$I = Aq\left(\frac{D_p}{L_p}p_{n0} + \frac{D_n}{L_n}n_{p0}\right)\left(e^{V/V_T} - 1\right)$$

Substituting for $p_{n0} = n_i^2/N_D$ and for $n_{p0} = n_i^2/N_A$ gives

$$I = Aqn_i^2\left(\frac{D_p}{L_pN_D} + \frac{D_n}{L_nN_A}\right)\left(e^{V/V_T} - 1\right) \tag{3.39}$$

From this equation we note that for a negative V (reverse bias) with a magnitude of a few times V_T (25.9 mV), the exponential term becomes essentially zero, and the current across the junction becomes negative and constant. From our qualitative description in Section 3.5.1, we know that this current must be I_S. Thus,

$$I = I_S\left(e^{V/V_T} - 1\right) \tag{3.40}$$

where

$$I_S = Aqn_i^2\left(\frac{D_p}{L_pN_D} + \frac{D_n}{L_nN_A}\right) \tag{3.41}$$

Figure 3.13 shows the I–V characteristic of the pn junction (Eq. 3.40). Observe that in the reverse direction the current saturates at a value equal to $-I_S$. For this reason, I_S is given the name **saturation current**. From Eq. (3.41) we see that I_S is directly proportional to the cross-sectional area A of the junction. Thus, another name for I_S, one we prefer to use in this book, is the junction **scale current**. Typical values for I_S, for junctions of various areas, range from 10^{-18} A to 10^{-12} A.

Besides being proportional to the junction area A, the expression for I_S in Eq. (3.41) indicates that I_S is proportional to n_i^2, which is a very strong function of temperature (see Eq. 3.2).

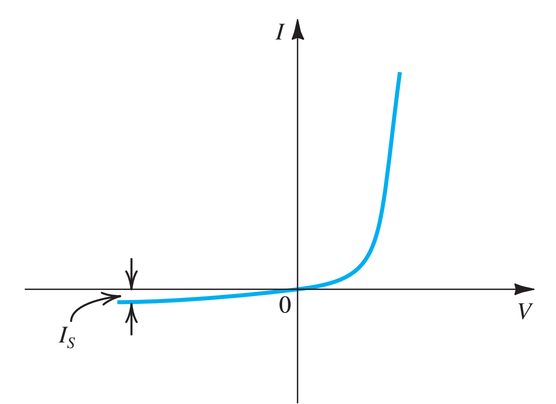


Figure 3.13 The pn junction I–V characteristic.

Example 3.6

For the *pn* junction considered in Example 3.5 for which $N_A = 10^{18}/\text{cm}^3$, $N_D = 10^{16}/\text{cm}^3$, $A = 10^{-4}\,\text{cm}^2$, and $n_i = 1.5 \times 10^{10}/\text{cm}^3$, let $L_p = 5\ \mu\text{m}$, $L_n = 10\ \mu\text{m}$, D_p (in the n region) $= 10\ \text{cm}^2/\text{V}\cdot\text{s}$, and D_n (in the p region) $= 18\ \text{cm}^2/\text{V}\cdot\text{s}$. The *pn* junction is forward biased and conducting a current $I = 0.1$ mA. Calculate: (a) I_S; (b) the forward-bias voltage V; and (c) the component of the current I due to hole injection and that due to electron injection across the junction.

Solution

(a) Using Eq. (3.41), we find I_S as

$$I_S = 10^{-4} \times 1.6 \times 10^{-19} \times \left(1.5 \times 10^{10}\right)^2 \times \left(\frac{10}{5 \times 10^{-4} \times 10^{16}} + \frac{18}{10 \times 10^{-4} \times 10^{18}}\right)$$
$$= 7.3 \times 10^{-15}\,\text{A}$$

(b) In the forward direction,

$$I = I_S\left(e^{V/V_T} - 1\right) \simeq I_S e^{V/V_T}$$

Thus,

$$V = V_T \ln\left(\frac{I}{I_S}\right)$$

For $I = 0.1$ mA,

$$V = 25.9 \times 10^{-3} \ln\left(\frac{0.1 \times 10^{-3}}{7.3 \times 10^{-15}}\right)$$
$$= 0.605\ \text{V}$$

(c) The hole-injection component of I can be found using Eq. (3.37)

$$I_p = Aq\frac{D_p}{L_p}p_{n0}\left(e^{V/V_T} - 1\right) = Aq\frac{D_p}{L_p}\frac{n_i^2}{N_D}\left(e^{V/V_T} - 1\right)$$

Similarly, I_n can be found using Eq. (3.39),

$$I_n = Aq\frac{D_n}{L_n}\frac{n_i^2}{N_A}\left(e^{V/V_T} - 1\right)$$

Thus,

$$\frac{I_p}{I_n} = \left(\frac{D_p}{D_n}\right)\left(\frac{L_n}{L_p}\right)\left(\frac{N_A}{N_D}\right)$$

For our case,

$$\frac{I_p}{I_n} = \frac{10}{18} \times \frac{10}{5} \times \frac{10^{18}}{10^{16}} = 1.11 \times 10^2 = 111$$

Example 3.6 *continued*

Thus most of the current is conducted by holes injected into the n region.

Specifically,

$$I_p = \frac{111}{112} \times 0.1 = 0.0991 \text{ mA}$$

$$I_n = \frac{1}{112} \times 0.1 = 0.0009 \text{ mA}$$

This stands to reason, since the p material has a doping concentration 100 times that of the n material.

EXERCISES

3.10 Show that if $N_A \gg N_D$,

$$I_S \simeq Aqn_i^2 \frac{D_p}{L_p N_D}$$

3.11 For the pn junction in Example 3.6, find the value of I_S and that of the current I at $V = 0.605$ V (same voltage found in Example 3.6 at a current $I = 0.1$ mA) if N_D is reduced by a factor of 2.
Ans. 1.46×10^{-14} A; 0.2 mA

3.12 For the pn junction considered in Examples 3.5 and 3.6, find the width of the depletion region W corresponding to the forward-bias voltage found in Example 3.6. (*Hint*: Use the formula in Eq. (3.31) with V_R replaced with $-V_F$.)
Ans. 0.166 μm

3.13 For the pn junction considered in Examples 3.5 and 3.6, find the width of the depletion region W and the charge stored in the depletion region Q_J when a 2-V reverse bias is applied. Also find the value of the reverse current I.
Ans. 0.608 μm; 9.63 pC; 7.3×10^{-15} A

3.5.3 Reverse Breakdown

The description of the operation of the pn junction in the reverse direction, and the $I-V$ relationship of the junction in Eq. (3.40), indicate that at a reverse-bias voltage $-V$, with $V \gg V_T$, the reverse current that flows across the junction is approximately equal to I_S and thus is very small. However, as the magnitude of the reverse-bias voltage V is increased, a value is reached at which a very large reverse current flows as shown in Fig. 3.14. Observe that as V reaches the value V_Z, the dramatic increase in reverse current is accompanied by a very small increase in the reverse voltage; that is, the reverse voltage across the junction

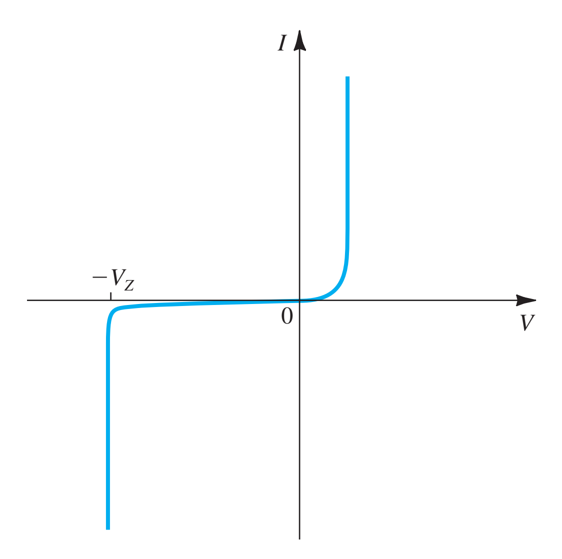


Figure 3.14 The *I–V* characteristic of the *pn* junction showing the rapid increase in reverse current in the breakdown region.

remains very close to the value V_Z. The phenomenon that occurs at $V = V_Z$ is known as **junction breakdown**. It is not a destructive phenomenon. That is, the *pn* junction can be repeatedly operated in the breakdown region without a permanent effect on its characteristics. This, however, is predicated on the assumption that the magnitude of the reverse-breakdown current is limited by the external circuit to a "safe" value. The "safe" value is one that results in the limitation of the power dissipated in the junction to a safe, allowable level.

There are two possible mechanisms for *pn* junction breakdown: the **zener effect**[7] and the **avalanche effect**. If a *pn* junction breaks down with a breakdown voltage $V_Z < 5$ V, the breakdown mechanism is usually the zener effect. Avalanche breakdown occurs when V_Z is greater than approximately 7 V. For junctions that break down between 5 V and 7 V, the breakdown mechanism can be either the zener or the avalanche effect or a combination of the two.

Zener breakdown occurs when the electric field in the depletion layer increases to the point of breaking covalent bonds and generating electron–hole pairs. The electrons generated in this way will be swept by the electric field into the *n* side and the holes into the *p* side. Thus these electrons and holes constitute a reverse current across the junction. Once the zener effect starts, a large number of carriers can be generated, with a negligible increase in the junction voltage. Thus the reverse current in the breakdown region will be large and its value must be determined by the external circuit, while the reverse voltage appearing between the diode terminals will remain close to the specified breakdown voltage V_Z.

The other breakdown mechanism, avalanche breakdown, occurs when the minority carriers that cross the depletion region under the influence of the electric field gain sufficient kinetic energy to be able to break covalent bonds in atoms with which they collide. The carriers liberated by this process may have sufficiently high energy to be able to cause other carriers to be liberated in another ionizing collision. This process keeps repeating in the fashion of an avalanche, with the result that many carriers are created that are able to support any value of

[7] Named after an early worker in the area. Note that the subscript *Z* in V_Z denotes *zener*. We will use V_Z to denote the breakdown voltage whether the breakdown mechanism is the zener effect or the avalanche effect.

reverse current, as determined by the external circuit, with a negligible change in the voltage drop across the junction.

As will be seen in Chapter 4, some *pn* junction diodes are fabricated to operate specifically in the breakdown region, where use is made of the nearly constant voltage V_Z.

3.6 Capacitive Effects in the *pn* Junction

There are two charge-storage mechanisms in the *pn* junction. One is associated with the charge stored in the depletion region, and the other is associated with the minority-carrier charge stored in the *n* and *p* materials as a result of the concentration profiles established by carrier injection. While the first is easier to see when the *pn* junction is reverse biased, the second is in effect only when the junction is forward biased.

3.6.1 Depletion or Junction Capacitance

When a *pn* junction is reverse biased with a voltage V_R, the charge stored on either side of the depletion region is given by Eq. (3.32),

$$Q_J = A\sqrt{2\epsilon_s q \frac{N_A N_D}{N_A + N_D}(V_0 + V_R)}$$

Thus, for a given *pn* junction,

$$Q_J = \alpha\sqrt{V_0 + V_R} \tag{3.42}$$

where α is given by

$$\alpha = A\sqrt{2\epsilon_s q \frac{N_A N_D}{N_A + N_D}} \tag{3.43}$$

Thus Q_J is nonlinearly related to V_R, as shown in Fig. 3.15. This nonlinear relationship makes it difficult to define a capacitance that accounts for the need to change Q_J whenever V_R is

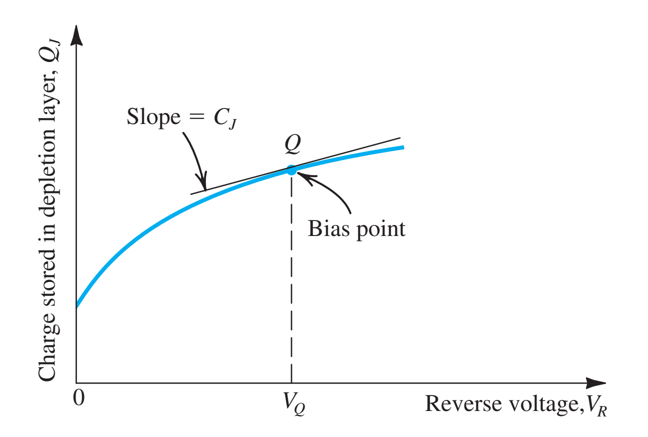


Figure 3.15 The charge stored on either side of the depletion layer as a function of the reverse voltage V_R.

changed. We can, however, assume that the junction is operating at a point such as Q, as indicated in Fig. 3.15, and define a capacitance C_j that relates the change in the charge Q_J to a change in the voltage V_R,

$$C_j = \left.\frac{dQ_J}{dV_R}\right|_{V_R=V_Q} \tag{3.44}$$

This incremental-capacitance approach turns out to be quite useful in electronic circuit design, as we shall see throughout this book.

Using Eq. (3.44) together with Eq. (3.42) yields

$$C_j = \frac{\alpha}{2\sqrt{V_0 + V_R}} \tag{3.45}$$

The value of C_j at zero reverse bias can be obtained from Eq. (3.45) as

$$C_{j0} = \frac{\alpha}{2\sqrt{V_0}} \tag{3.46}$$

which enables us to express C_j as

$$C_j = \frac{C_{j0}}{\sqrt{1 + \frac{V_R}{V_0}}} \tag{3.47}$$

where C_{j0} is given by Eq. (3.46) or alternatively if we substitute for α from Eq. (3.43) by

$$C_{j0} = A\sqrt{\left(\frac{\epsilon_s q}{2}\right)\left(\frac{N_A N_D}{N_A + N_D}\right)\left(\frac{1}{V_0}\right)} \tag{3.48}$$

Before leaving the subject of depletion-region or junction capacitance we point out that in the *pn* junction we have been studying, the doping concentration is made to change abruptly at the junction boundary. Such a junction is known as an **abrupt junction**. There is another type of *pn* junction in which the carrier concentration is made to change gradually from one side of the junction to the other. To allow for such a **graded junction**, the formula for the junction capacitance (Eq. 3.47) can be written in the more general form

$$C_j = \frac{C_{j0}}{\left(1 + \frac{V_R}{V_0}\right)^m} \tag{3.49}$$

where m is a constant called the **grading coefficient**, whose value ranges from 1/3 to 1/2 depending on the manner in which the concentration changes from the p to the n side.

EXERCISE

3.14 For the *pn* junction considered in Examples 3.5 and 3.6, find C_{j0} and C_j at $V_R = 2$ V. Recall that $V_0 = 0.814$ V, $N_A = 10^{18}/\text{cm}^3$, $N_D = 10^{16}/\text{cm}^3$, and $A = 10^{-4}\ \text{cm}^2$.
Ans. 3.2 pF; 1.7 pF

3.6.2 Diffusion Capacitance

Consider a forward-biased *pn* junction. In steady state, minority-carrier distributions in the *p* and *n* materials are established, as shown in Fig. 3.12. Thus a certain amount of excess minority-carrier charge is stored in each of the *p* and *n* bulk regions (outside the depletion region). If the terminal voltage *V* changes, this charge will have to change before a new steady state is achieved. This charge-storage phenomenon gives rise to another capacitive effect, distinctly different from that due to charge storage in the depletion region.

To calculate the excess minority-carrier charge, refer to Fig. 3.12. The excess hole charge stored in the *n* region can be found from the shaded area under the exponential as follows:[8]

$$\begin{aligned} Q_p &= Aq \times \text{shaded area under the } p_n(x) \text{curve} \\ &= Aq[p_n(x_n) - p_{n0}]L_p \end{aligned}$$

Substituting for $p_n(x_n)$ from Eq. (3.33) and using Eq. (3.37) enables us to express Q_p as

$$Q_p = \frac{L_p^2}{D_p} I_p \tag{3.50}$$

The factor $\left(L_p^2/D_p\right)$ that relates Q_p to I_p is a useful device parameter that has the dimension of time (s) and is denoted τ_p

$$\tau_p = \frac{L_p^2}{D_p} \tag{3.51}$$

Thus,

$$Q_p = \tau_p I_p \tag{3.52}$$

The time constant τ_p is known as the excess **minority-carrier (hole) lifetime**. *It is the average time it takes for a hole injected into the n region to recombine with a majority electron.* This definition of τ_p implies that the entire charge Q_p disappears and has to be replenished every τ_p seconds. The current that accomplishes the replenishing is $I_p = Q_p/\tau_p$. This is an alternate derivation for Eq. (3.52).

---

[8]Recall that the area under an exponential curve $Ae^{-x/B}$ is equal to AB.

A relationship similar to that in Eq. (3.52) can be developed for the electron charge stored in the p region,

$$Q_n = \tau_n I_n \tag{3.53}$$

where τ_n is the electron lifetime in the p region. The total excess minority-carrier charge can be obtained by adding together Q_p and Q_n,

$$Q = \tau_p I_p + \tau_n I_n \tag{3.54}$$

This charge can be expressed in terms of the diode current $I = I_p + I_n$ as

$$Q = \tau_T I \tag{3.55}$$

where τ_T is called the **mean transit time** of the junction. Obviously, τ_T is related to τ_p and τ_n. Furthermore, for most practical devices, one side of the junction is much more heavily doped than the other. For instance, if $N_A \gg N_D$, one can show that $I_p \gg I_n$, $I \simeq I_p$, $Q_p \gg Q_n$, $Q \simeq Q_p$, and thus $\tau_T \simeq \tau_p$.

For small changes around a bias point, we can define an **incremental diffusion capacitance** C_d as

$$C_d = \frac{dQ}{dV} \tag{3.56}$$

and can show that

$$C_d = \left(\frac{\tau_T}{V_T}\right) I \tag{3.57}$$

where I is the forward-bias current. Note that C_d is directly proportional to the forward current I and thus is negligibly small when the diode is reverse biased. Also note that to keep C_d small, the transit time τ_T must be made small, an important requirement for a *pn* junction intended for high-speed or high-frequency operation.

EXERCISES

3.15 Use the definition of C_d in Eq. (3.56) to derive the expression in Eq. (3.57) by means of Eqs. (3.55) and (3.40).

3.16 For the *pn* junction considered in Examples 3.5 and 3.6 for which $D_p = 10\ \text{cm}^2/\text{V}\cdot\text{s}$, and $L_p = 5\ \mu\text{m}$, find τ_p and C_d at a forward-bias current of 0.1 mA. Recall that for this junction, $I_p \simeq I$.
Ans. 25 ns; 96.5 pF

Summary

- Today's microelectronics technology is almost entirely based on the semiconductor material silicon. If a circuit is to be fabricated as a monolithic integrated circuit (IC) it is made using a single silicon crystal, no matter how large the circuit is (a recent chip contains 4.31 billion transistors).
- In a crystal of intrinsic or pure silicon, the atoms are held in position by covalent bonds. At very low temperatures, all the bonds are intact, and no charge carriers are available to conduct electrical current. Thus, at such low temperatures, silicon behaves as an insulator.
- At room temperature, thermal energy causes some of the covalent bonds to break, thus generating free electrons and holes that become available for current conduction.
- Current in semiconductors is carried by free electrons and holes. Their numbers are equal and relatively small in intrinsic silicon.
- The conductivity of silicon can be increased dramatically by introducing small amounts of appropriate impurity materials into the silicon crystal in a process called doping.
- There are two kinds of doped semiconductor: n-type, in which electrons are abundant, and p-type, in which holes are abundant.
- There are two mechanisms for the transport of charge carriers in semiconductors: drift and diffusion.
- Carrier drift results when an electric field E is applied across a piece of silicon. The electric field accelerates the holes in the direction of E and the electrons in the direction opposite to E. These two current components add together to produce a drift current in the direction of E.
- Carrier diffusion occurs when the concentration of charge carriers is made higher in one part of the silicon crystal than in other parts. To establish a steady-state diffusion current, a carrier concentration gradient must be maintained in the silicon crystal.
- A basic semiconductor structure is the pn junction. It is fabricated in a silicon crystal by creating a p region in close proximity to an n region. The pn junction is a diode and plays a dominant role in the structure and operation of transistors.
- When the terminals of the pn junction are left open, no current flows externally. However, two equal and opposite currents, I_D and I_S, flow across the junction, and equilibrium is maintained by a built-in voltage V_0 that develops across the junction, with the n side positive relative to the p side. Note, however, that the voltage across an open junction is 0 V, since V_0 is canceled by potentials appearing at the metal-to-semiconductor connection interfaces.
- The voltage V_0 appears across the depletion region, which extends on both sides of the junction.
- The diffusion current I_D is carried by holes diffusing from p to n and electrons diffusing from n to p. I_D flows from p to n, which is the forward direction of the junction. Its value depends on V_0.
- The drift current I_S is carried by thermally generated minority electrons in the p material that are swept across the depletion layer into the n side, and by thermally generated minority holes in the n side that are swept across the depletion region into the p side. I_S flows from n to p, in the reverse direction of the junction, and its value is a strong function of temperature but independent of V_0.
- Forward biasing the pn junction, that is, applying an external voltage V that makes p more positive than n, reduces the barrier voltage to $V_0 - V$ and results in an exponential increase in I_D while I_S remains unchanged. The net result is a substantial current $I = I_D - I_S$ that flows across the junction and through the external circuit.
- Applying a negative V reverse biases the junction and increases the barrier voltage, with the result that I_D is reduced to almost zero and the net current across the junction becomes the very small reverse current I_S.
- If the reverse voltage is increased in magnitude to a value V_Z specific to the particular junction, the junction breaks down, and a large reverse current flows. The value of the reverse current must be limited by the external circuit.
- Whenever the voltage across a pn junction is changed, some time has to pass before steady state is reached. This is due to the charge-storage effects in the junction, which are modeled by two capacitances: the junction capacitance C_j and the diffusion capacitance C_d.
- For future reference, we present in Table 3.1 a summary of pertinent relationships and the values of physical constants.

Table 3.1 Summary of Important Equations

Quantity	Relationship	Values of Constants and Parameters (for Intrinsic Si at $T = 300$ K)
Carrier concentration in intrinsic silicon (cm^{-3})	$n_i = BT^{3/2}e^{-E_g/2kT}$	$B = 7.3 \times 10^{15}\ cm^{-3}K^{-3/2}$ $E_g = 1.12$ eV $k = 8.62 \times 10^{-5}$ eV/K $n_i = 1.5 \times 10^{10}/cm^3$
Diffusion current density (A/cm^2)	$J_p = -qD_p \frac{dp}{dx}$ $J_n = qD_n \frac{dn}{dx}$	$q = 1.60 \times 10^{-19}$ coulomb $D_p = 12\ cm^2/s$ $D_n = 34\ cm^2/s$
Drift current density (A/cm^2)	$J_{\text{drift}} = q(p\mu_p + n\mu_n)E$	$\mu_p = 480\ cm^2/V \cdot s$ $\mu_n = 1350\ cm^2/V \cdot s$
Resistivity ($\Omega \cdot cm$)	$\rho = 1/[q(p\mu_p + n\mu_n)]$	μ_p and μ_n decrease with the increase in doping concentration
Relationship between mobility and diffusivity	$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = V_T$	$V_T = kT/q \simeq 25.9$ mV
Carrier concentration in *n*-type silicon (cm^{-3})	$n_{n0} \simeq N_D$ $p_{n0} = n_i^2/N_D$	
Carrier concentration in *p*-type silicon (cm^{-3})	$p_{p0} \simeq N_A$ $n_{p0} = n_i^2/N_A$	
Junction built-in voltage (V)	$V_0 = V_T \ln\left(\frac{N_A N_D}{n_i^2}\right)$	
Width of depletion region (cm)	$\frac{x_n}{x_p} = \frac{N_A}{N_D}$ $W = x_n + x_p$ $= \sqrt{\frac{2\epsilon_s}{q}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)(V_0 + V_R)}$	$\epsilon_s = 11.7\epsilon_0$ $\epsilon_0 = 8.854 \times 10^{-14}$ F/cm

Table 3.1 *continued*

Quantity	Relationship	Values of Constants and Parameters (for Intrinsic Si at $T = 300$ K)
Charge stored in depletion layer (coulomb)	$Q_J = q\dfrac{N_A N_D}{N_A + N_D} AW$	
Forward current (A)	$I = I_p + I_n$ $I_p = Aqn_i^2 \dfrac{D_p}{L_p N_D}\left(e^{V/V_T} - 1\right)$ $I_n = Aqn_i^2 \dfrac{D_n}{L_n N_A}\left(e^{V/V_T} - 1\right)$	
Saturation current (A)	$I_S = Aqn_i^2\left(\dfrac{D_p}{L_p N_D} + \dfrac{D_n}{L_n N_A}\right)$	
I–V relationship	$I = I_S\left(e^{V/V_T} - 1\right)$	
Minority-carrier lifetime (s)	$\tau_p = L_p^2/D_p \qquad \tau_n = L_n^2/D_n$	$L_p, L_n = 1\ \mu\text{m to } 100\ \mu\text{m}$ $\tau_p, \tau_n = 1\ \text{ns to } 10^4\ \text{ns}$
Minority-carrier charge storage (coulomb)	$Q_p = \tau_p I_p \qquad Q_n = \tau_n I_n$ $Q = Q_p + Q_n = \tau_T I$	
Depletion capacitance (F)	$C_{j0} = A\sqrt{\left(\dfrac{\epsilon_s q}{2}\right)\left(\dfrac{N_A N_D}{N_A + N_D}\right)\dfrac{1}{V_0}}$ $C_j = C_{j0} \Big/ \left(1 + \dfrac{V_R}{V_0}\right)^m$	$m = \dfrac{1}{3}\ \text{to}\ \dfrac{1}{2}$
Diffusion capacitance (F)	$C_d = \left(\dfrac{\tau_T}{V_T}\right) I$	

If in the following problems the need arises for the values of particular parameters or physical constants that are not stated, please consult Table 3.1.

Section 3.1: Intrinsic Semiconductors

3.1 Find values of the intrinsic carrier concentration n_i for silicon at −55°C, 0°C, 20°C, 75°C, and 125°C. At each temperature, what fraction of the atoms is ionized? Recall that a silicon crystal has approximately 5×10^{22} atoms/cm^3.

3.2 Calculate the value of n_i for gallium arsenide (GaAs) at $T = 300$ K. The constant $B = 3.56 \times 10^{14}$ cm^{-3}K$^{-3/2}$ and the bandgap voltage $E_g = 1.42$ eV.

Section 3.2: Doped Semiconductors

3.3 For a p-type silicon in which the dopant concentration $N_A = 5 \times 10^{18}$/cm^3, find the hole and electron concentrations at $T = 300$ K.

3.4 For a silicon crystal doped with phosphorus, what must N_D be if at $T = 300$ K the hole concentration drops below the intrinsic level by a factor of 10^8?

3.5 In a phosphorus-doped silicon layer with impurity concentration of 10^{17}/cm^3, find the hole and electron concentrations at 27°C and 125°C.

Section 3.3: Current Flow in Semiconductors

3.6 A young designer, aiming to develop intuition concerning conducting paths within an integrated circuit, examines the end-to-end resistance of a connecting bar 10-μm long, 3-μm wide, and 1 μm thick, made of various materials. The designer considers:

(a) intrinsic silicon
(b) n-doped silicon with $N_D = 5 \times 10^{16}$/cm^3
(c) n-doped silicon with $N_D = 5 \times 10^{18}$/cm^3
(d) p-doped silicon with $N_A = 5 \times 10^{16}$/cm^3
(e) aluminum with resistivity of 2.8 μΩ·cm

Find the resistance in each case. For intrinsic silicon, use the data in Table 3.1. For doped silicon, assume $\mu_n = 3\mu_p = 1200$ cm^2/V·s. (Recall that $R = \rho L/A$.)

3.7 Contrast the electron and hole drift velocities through a 10-μm layer of intrinsic silicon across which a voltage of 3 V is imposed. Let $\mu_n = 1350$ cm^2/V·s and $\mu_p = 480$ cm^2/V·s.

3.8 Find the current that flows in a silicon bar of 10-μm length having a 5-μm × 4-μm cross-section and having free-electron and hole densities of 10^4/cm^3 and 10^{16}/cm^3, respectively, when a 1 V is applied end-to-end. Use $\mu_n = 1200$ cm^2/V·s and $\mu_p = 500$ cm^2/V·s.

3.9 In a 10-μm-long bar of donor-doped silicon, what donor concentration is needed to realize a current density of 2 mA/μm^2 in response to an applied voltage of 1 V? (*Note*: Although the carrier mobilities change with doping concentration, as a first approximation you may assume μ_n to be constant and use 1350 cm^2/V·s, the value for intrinsic silicon.)

3.10 Holes are being steadily injected into a region of n-type silicon (connected to other devices, the details of which are not important for this question). In the steady state, the excess-hole concentration profile shown in Fig. P3.10 is established in the n-type silicon region. Here "excess" means over and above the thermal-equilibrium concentration (in the absence of hole injection), denoted p_{n0}. If $N_D = 10^{16}$/cm^3, $n_i = 1.5 \times 10^{10}$/cm^3, $D_p = 12$ cm^2/s, and $W = 50$ nm, find the density of the current that will flow in the x direction.

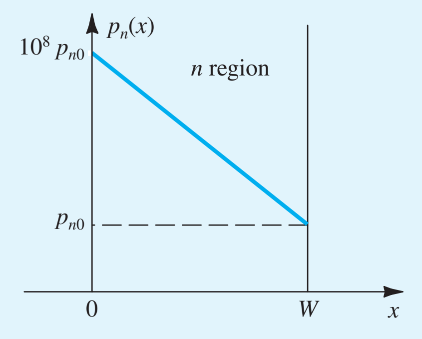


Figure P3.10

3.11 Both the carrier mobility and the diffusivity decrease as the doping concentration of silicon is increased. Table P3.11 provides a few data points for μ_n and μ_p versus doping concentration. Use the Einstein relationship to obtain the corresponding values for D_n and D_p.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

Table P3.11

Doping Concentration (carriers/cm^3)	μ_n (cm^2/V · s)	μ_p (cm^2/V · s)	D_n (cm^2/s)	D_p (cm^2/s)
Intrinsic	1350	480		
10^{16}	1200	400		
10^{17}	750	260		
10^{18}	380	160		

Section 3.4: The *pn* Junction

3.12 Calculate the built-in voltage of a junction in which the p and n regions are doped equally with 5×10^{16} atoms/cm^3. Assume $n_i = 1.5 \times 10^{10}$/cm^3. With the terminals left open, what is the width of the depletion region, and how far does it extend into the p and n regions? If the cross-sectional area of the junction is 20 μm^2, find the magnitude of the charge stored on either side of the junction.

3.13 If, for a particular junction, the acceptor concentration is 10^{17}/cm^3 and the donor concentration is 10^{16}/cm^3, find the junction built-in voltage. Assume $n_i = 1.5 \times 10^{10}$/cm^3. Also, find the width of the depletion region (W) and its extent in each of the p and n regions when the junction terminals are left open. Calculate the magnitude of the charge stored on either side of the junction. Assume that the junction area is 100 μm^2.

3.14 Estimate the total charge stored in a 0.1-μm depletion layer on one side of a 10-μm × 10-μm junction. The doping concentration on that side of the junction is 10^{18}/cm^3.

3.15 In a *pn* junction for which $N_A \gg N_D$, and the depletion layer exists mostly on the shallowly doped side with $W =$ 0.2 μm, find V_0 if $N_D = 10^{16}$/cm^3. Also calculate Q_J for the case $A = 10$ μm^2.

3.16 By how much does V_0 change if N_A or N_D is increased by a factor of 10?

Section 3.5: The *pn* Junction with an Applied Voltage

3.17 If a 3-V reverse-bias voltage is applied across the junction specified in Problem 3.13, find W and Q_J.

3.18 Show that for a *pn* junction reverse-biased with a voltage V_R, the depletion-layer width W and the charge stored on either side of the junction, Q_J, can be expressed as

$$W = W_0\sqrt{1+\frac{V_R}{V_0}}$$

$$Q_J = Q_{J0}\sqrt{1+\frac{V_R}{V_0}}$$

where W_0 and Q_{J0} are the values in equilibrium.

3.19 In a forward-biased *pn* junction show that the ratio of the current component due to hole injection across the junction to the component due to electron injection is given by

$$\frac{I_p}{I_n} = \frac{D_p}{D_n}\frac{L_n}{L_p}\frac{N_A}{N_D}$$

Evaluate this ratio for the case $N_A = 10^{18}$/cm^3, $N_D = 10^{16}$/cm^3, $L_p = 5$ μm, $L_n = 10$ μm, $D_p = 10$ cm^2/s, and $D_n = 20$ cm^2/s, and hence find I_p and I_n for the case in which the *pn* junction is conducting a forward current $I = 100$ μA.

3.20 Calculate I_S and the current I for $V = 750$ mV for a *pn* junction for which $N_A = 10^{17}$/cm^3, $N_D = 10^{16}$/cm^3, $A = 100$ μm^2, $n_i = 1.5 \times 10^{10}$/cm^3, $L_p = 5$ μm, $L_n = 10$ μm, $D_p = 10$ cm^2/s, and $D_n = 18$ cm^2/s.

3.21 Assuming that the temperature dependence of I_S arises mostly because I_S is proportional to n_i^2, use the expression for n_i in Eq. (3.2) to determine the factor by which n_i^2 changes as T changes from 300 K to 305 K. This will be approximately the same factor by which I_S changes for a 5°C rise in temperature. What is the factor?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

3.22 A p^+n junction is one in which the doping concentration in the p region is much greater than that in the n region. In such a junction, the forward current is mostly due to hole injection across the junction. Show that

$$I \simeq I_p = Aqn_i^2 \frac{D_p}{L_p N_D}\left(e^{V/V_T} - 1\right)$$

For the specific case in which $N_D = 10^{17}/\text{cm}^3$, $D_p = 10\ \text{cm}^2/\text{s}$, $L_p = 10\ \mu\text{m}$, and $A = 10^4\ \mu\text{m}^2$, find I_S and the voltage V obtained when $I = 1$ mA. Assume operation at 300 K where $n_i = 1.5 \times 10^{10}/\text{cm}^3$.

3.23 A pn junction for which the breakdown voltage is 12 V has a rated (i.e., maximum allowable) power dissipation of 0.25 W. What continuous current in the breakdown region will raise the dissipation to half the rated value? If breakdown occurs for only 10 ms in every 20 ms, what average breakdown current is allowed?

Section 3.6: Capacitive Effects in the *pn* Junction

3.24 For the pn junction specified in Problem 3.13, find C_{j0} and C_j at $V_R = 3$ V.

3.25 For a particular junction for which $C_{j0} = 0.4$ pF, $V_0 = 0.75$ V, and $m = 1/3$, find C_j at reverse-bias voltages of 1 V and 10 V.

3.26 The junction capacitance C_j can be thought of as that of a parallel-plate capacitor and thus given by

$$C_j = \frac{\epsilon A}{W}$$

Show that this approach leads to a formula identical to that obtained by combining Eqs. (3.43) and (3.45) [or equivalently, by combining Eqs. (3.47) and (3.48)].

3.27 A pn junction operating in the forward-bias region with a current I of 1 mA is found to have a diffusion capacitance of 5 pF. What diffusion capacitance do you expect this junction to have at $I = 0.1$ mA? What is the mean transit time for this junction?

3.28 For the p^+n junction specified in Problem 3.22, find τ_p and calculate the excess minority-carrier charge and the value of the diffusion capacitance at $I = 0.1$ mA.

***3.29** A **short-base diode** is one where the widths of the p and n regions are much smaller than L_n and L_p, respectively. As a result, the excess minority-carrier distribution in each region is a straight line rather than the exponentials shown in Fig. 3.12.

(a) For the short-base diode, sketch a figure corresponding to Fig. 3.12 and assume as in Fig. 3.12 that $N_A \gg N_D$.

(b) Following a derivation similar to that given in Section 3.5.2, show that if the widths of the p and n regions are denoted W_p and W_n then

$$I = Aqn_i^2\left[\frac{D_p}{(W_n - x_n)N_D} + \frac{D_n}{(W_p - x_p)N_A}\right]\left(e^{V/V_T} - 1\right)$$

and

$$Q_p = \frac{1}{2}\frac{(W_n - x_n)^2}{D_p} I_p$$

$$\simeq \frac{1}{2}\frac{W_n^2}{D_p} I_p, \text{ for } W_n \gg x_n$$

(c) Also, assuming $Q \simeq Q_p$, $I \simeq I_p$, show that

$$C_d = \frac{\tau_T}{V_T} I$$

where

$$\tau_T = \frac{1}{2}\frac{W_n^2}{D_p}$$

(d) If a designer wishes to limit C_d to 8 pF at $I = 1$ mA, what should W_n be? Assume $D_p = 10\ \text{cm}^2/\text{s}$.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

CHAPTER 4

Diodes

IN THIS CHAPTER YOU WILL LEARN

1. The characteristics of the ideal diode and how to analyze and design circuits containing multiple ideal diodes together with resistors and dc sources to realize useful and interesting nonlinear functions.
2. The details of the i–v characteristic of the junction diode (which was derived in Chapter 3) and how to use it to analyze diode circuits operating in the various bias regions: forward, reverse, and breakdown.
3. A simple but effective model of the diode i–v characteristic in the forward direction: the constant-voltage-drop model.
4. A powerful technique for the application and modeling of the diode (and in later chapters, transistors): dc-biasing the diode and modeling its operation for small signals around the dc operating point by means of the small-signal model.
5. The use of a string of forward-biased diodes and of diodes operating in the breakdown region (zener diodes), to provide constant dc voltages (voltage regulators).
6. Application of the diode in the design of rectifier circuits, which convert ac voltages to dc as needed for powering electronic equipment.
7. A number of other practical and important applications of diodes.

Introduction

In Chapters 1 and 2 we dealt almost entirely with linear circuits; any nonlinearity, such as that introduced by amplifier output saturation, was treated as a problem to be solved by the circuit designer. However, there are many other signal-processing functions that can be implemented only by nonlinear circuits. Examples include the generation of dc voltages from the ac power supply, and the generation of signals of various waveforms (e.g., sinusoids, square waves, pulses). Also, digital logic and memory circuits constitute a special class of nonlinear circuits.

The simplest and most fundamental nonlinear circuit element is the diode. Just like a resistor, the diode has two terminals; but unlike the resistor, which has a linear (straight-line) relationship between the current flowing through it and the voltage appearing across it, the diode has a nonlinear i–v characteristic.

This chapter is concerned with the study of diodes. In order to understand the essence of the diode function, we begin with a fictitious element, the ideal diode. We then introduce the silicon junction diode, explain its terminal characteristics, and provide techniques for the

analysis of diode circuits. The latter task involves the important subject of device modeling. Our study of modeling the diode characteristics will lay the foundation for our study of modeling transistor operation in the next three chapters.

Of the many applications of diodes, their use in the design of rectifiers (which convert ac to dc) is the most common. Therefore we shall study rectifier circuits in some detail and briefly look at a number of other diode applications. Further nonlinear circuits that utilize diodes and other devices will be found throughout the book, but particularly in Chapter 18.

The junction diode is nothing more than the *pn* junction we studied in Chapter 3, and most of this chapter is concerned with the study of silicon *pn*-junction diodes. In the last section, however, we briefly consider some specialized diode types, including the photodiode and the light-emitting diode.

4.1 The Ideal Diode

4.1.1 Current–Voltage Characteristic

The ideal diode may be considered to be the most fundamental nonlinear circuit element. It is a two-terminal device having the circuit symbol of Fig. 4.1(a) and the i–v characteristic shown in Fig. 4.1(b). The terminal characteristic of the ideal diode can be interpreted as follows: If a negative voltage (relative to the reference direction indicated in Fig. 4.1a) is applied to the diode, no current flows and the diode behaves as an open circuit (Fig. 4.1c). Diodes operated in this mode are said to be **reverse biased**, or operated in the reverse direction. An ideal diode has zero current when operated in the reverse direction and is said to be **cut off**, or simply **off**.

On the other hand, if a positive current (relative to the reference direction indicated in Fig. 4.1(a) is applied to the ideal diode, zero voltage drop appears across the diode. In other words, the ideal diode behaves as a short circuit in the *forward* direction (Fig. 4.1d); it passes any current with zero voltage drop. A **forward-biased** diode is said to be **turned on**, or simply **on**.

From the above description it should be noted that the external circuit must be designed to limit the forward current through a conducting diode, and the reverse voltage across a cutoff diode, to predetermined values. Figure 4.2 shows two diode circuits that illustrate this point. In the circuit of Fig. 4.2(a) the diode is obviously conducting. Thus its voltage drop will be zero, and the current through it will be determined by the +10-V supply and the 1-kΩ resistor as 10 mA. The diode in the circuit of Fig. 4.2(b) is obviously cut off, and thus its current will be zero, which in turn means that the entire 10-V supply will appear as reverse bias across the diode.

The positive terminal of the diode is called the **anode** and the negative terminal the **cathode**, a carryover from the days of vacuum-tube diodes. The i–v characteristic of the ideal diode (conducting in one direction and not in the other) should explain the choice of its arrow-like circuit symbol.

As should be evident from the preceding description, the i–v characteristic of the ideal diode is highly nonlinear; although it consists of two straight-line segments, they are at 90° to one another. A nonlinear curve that consists of straight-line segments is said to be **piecewise linear**. If a device having a piecewise-linear characteristic is used in a particular application in such a way that the signal across its terminals swings along only one of the linear segments, then the device can be considered a linear circuit element as far as that particular circuit

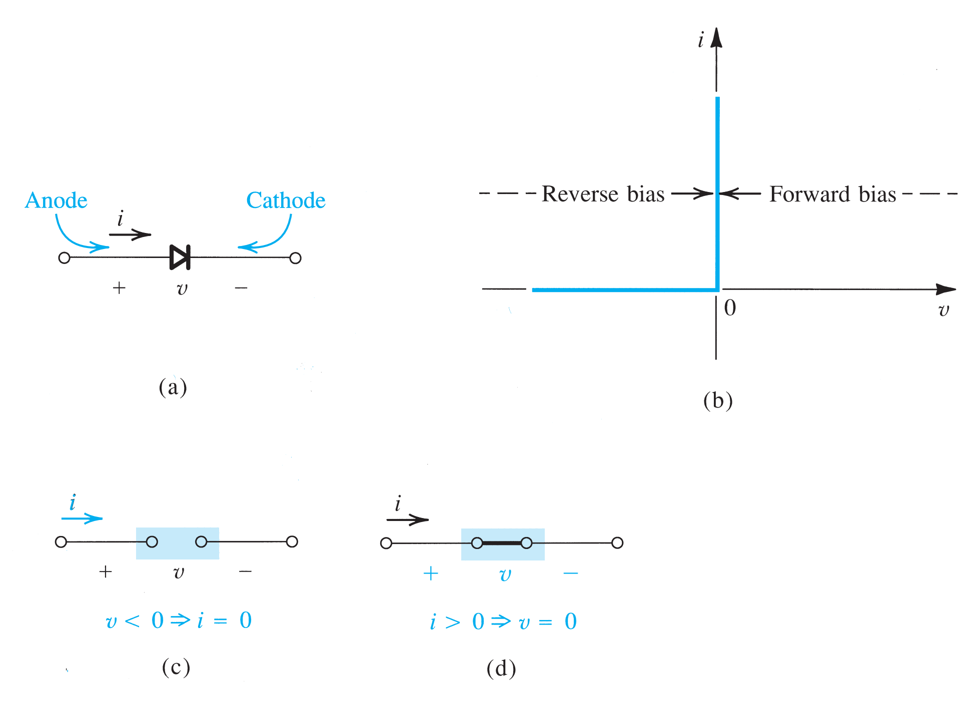


Figure 4.1 The ideal diode: **(a)** diode circuit symbol; **(b)** i–v characteristic; **(c)** equivalent circuit in the reverse direction; **(d)** equivalent circuit in the forward direction.

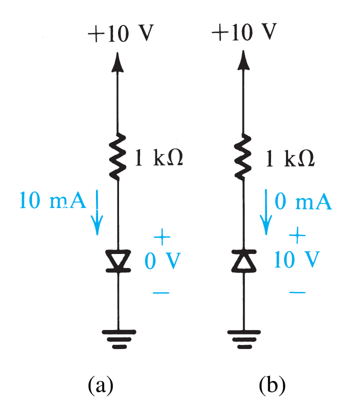


Figure 4.2 The two modes of operation of ideal diodes and the use of an external circuit to limit **(a)** the forward current and **(b)** the reverse voltage.

application is concerned. On the other hand, if signals swing past one or more of the break points in the characteristic, linear analysis is no longer possible.

4.1.2 A Simple Application: The Rectifier

A fundamental application of the diode, one that makes use of its severely nonlinear i–v curve, is the rectifier circuit shown in Fig. 4.3(a). The circuit consists of the series connection of a diode D and a resistor R. Let the input voltage v_I be the sinusoid shown in Fig. 4.3(b), and assume the diode to be ideal. During the positive half-cycles of the input sinusoid, the positive

v_I will cause current to flow through the diode in its forward direction. It follows that the diode voltage v_D will be very small—ideally zero. Thus the circuit will have the equivalent shown in Fig. 4.3(c), and the output voltage v_O will be equal to the input voltage v_I. On the other hand, during the negative half-cycles of v_I, the diode will not conduct. Thus the circuit will have the equivalent shown in Fig. 4.3(d), and v_O will be zero. Thus the output voltage will have the waveform shown in Fig. 4.3(e). Note that while v_I alternates in polarity and has a zero average value, v_O is unidirectional and has a finite average value or a *dc component*. Thus the circuit of Fig. 4.3(a) **rectifies** the signal and hence is called a **rectifier**. It can be used to generate dc from ac. We will study rectifier circuits in Section 4.5.

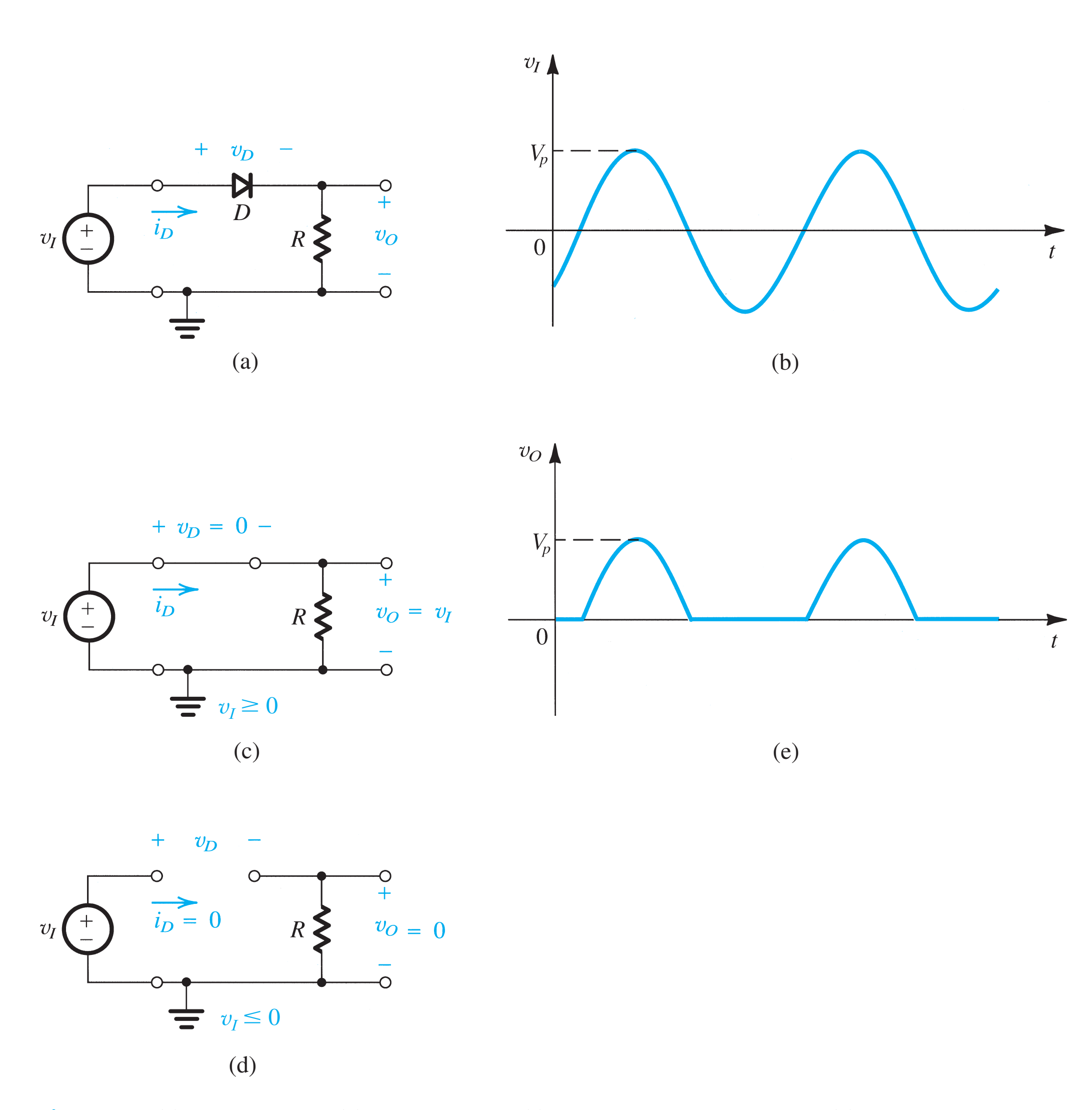


Figure 4.3 **(a)** Rectifier circuit. **(b)** Input waveform. **(c)** Equivalent circuit when $v_I \geq 0$. **(d)** Equivalent circuit when $v_I \leq 0$. **(e)** Output waveform.

EXERCISES

4.1 For the circuit in Fig. 4.3(a), sketch the transfer characteristic v_O versus v_I.
Ans. See Fig. E4.1

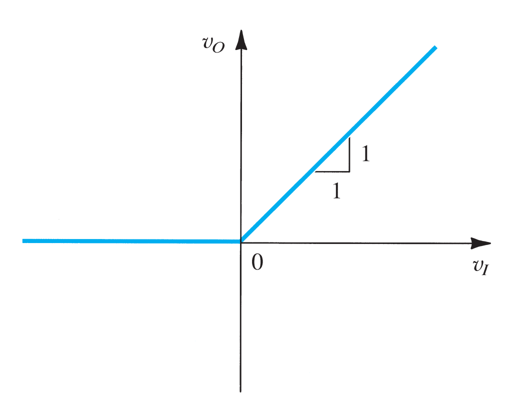


Figure E4.1

4.2 For the circuit in Fig. 4.3(a), sketch the waveform of v_D.
Ans. $v_D = v_I - v_O$, resulting in the waveform in Fig. E4.2

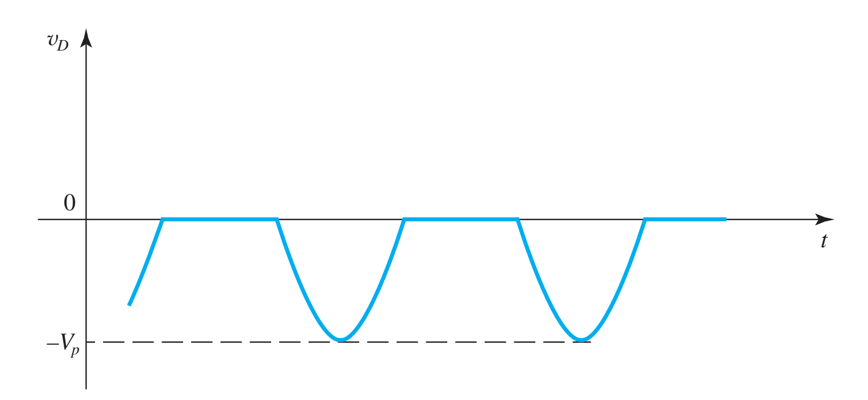


Figure E4.2

4.3 In the circuit of Fig. 4.3(a), let v_I have a peak value of 10 V and $R = 1\ \text{k}\Omega$. Find the peak value of i_D and the dc component of v_O. (*Hint*: The average value of half-sine waves is V_p/π.)
Ans. 10 mA; 3.18 V

Example 4.1

Figure 4.4(a) shows a circuit for charging a 12-V battery. If v_S is a sinusoid with 24-V peak amplitude, find the fraction of each cycle during which the diode conducts. Also, find the peak value of the diode current and the maximum reverse-bias voltage that appears across the diode.

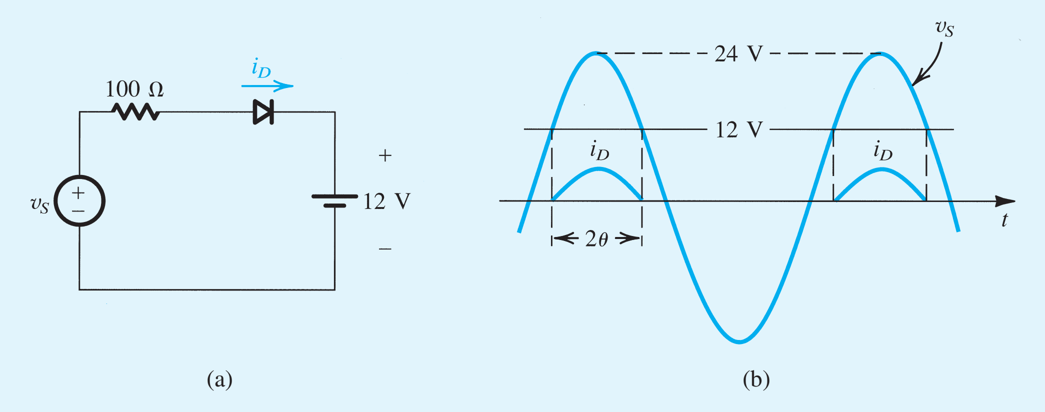


Figure 4.4 Circuit and waveforms for Example 4.1.

Solution

The diode conducts when v_S exceeds 12 V, as shown in Fig. 4.4(b). The conduction angle is 2θ, where θ is given by

$$24\cos\theta = 12$$

Thus $\theta = 60°$ and the conduction angle is 120°, or one-third of a cycle.

The peak value of the diode current is given by

$$I_d = \frac{24-12}{100} = 0.12 \text{ A}$$

The maximum reverse voltage across the diode occurs when v_S is at its negative peak and is equal to $24 + 12 = 36$ V.

4.1.3 Another Application: Diode Logic Gates

Diodes together with resistors can be used to implement digital logic functions. Figure 4.5 shows two diode logic gates. To see how these circuits function, consider a positive-logic system in which voltage values close to 0 V correspond to logic 0 (or low) and voltage values close to +5 V correspond to logic 1 (or high). The circuit in Fig. 4.5(a) has three inputs, v_A, v_B, and v_C. It is easy to see that diodes connected to +5-V inputs will conduct, thus clamping the output v_Y to a value equal to +5 V. This positive voltage at the output will keep the diodes whose inputs are low (around 0 V) cut off. Thus *the output will be high if one or more of the inputs are high*. The circuit therefore implements the **logic OR function**, which in Boolean

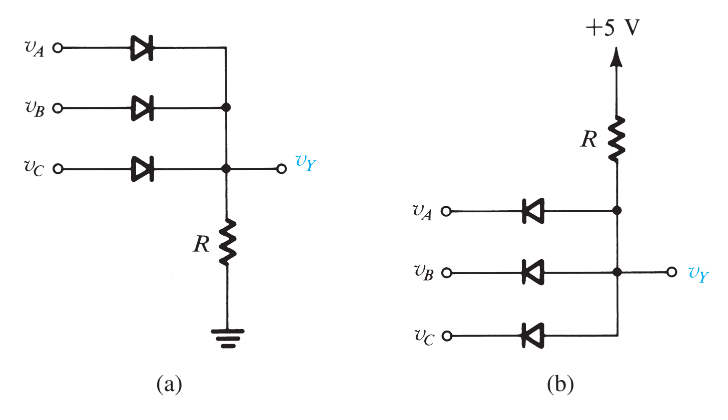


Figure 4.5 Diode logic gates: **(a)** OR gate; **(b)** AND gate (in a positive-logic system).

notation is expressed as

$$Y = A + B + C$$

Similarly, the reader is encouraged to show that using the same logic system mentioned above, the circuit of Fig. 4.5(b) implements the **logic AND function**,

$$Y = A \cdot B \cdot C$$

Example 4.2

Assuming the diodes to be ideal, find the values of I and V in the circuits of Fig. 4.6.

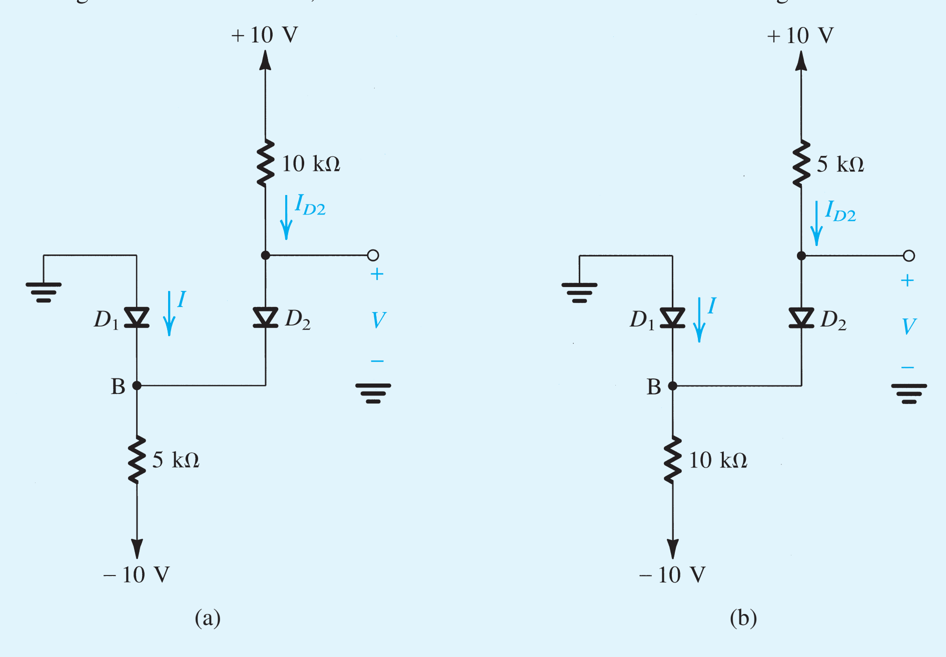


Figure 4.6 Circuits for Example 4.2.

Example 4.2 *continued*

Solution

In these circuits it might not be obvious at first sight whether none, one, or both diodes are conducting. In such a case, *we make a plausible assumption, proceed with the analysis, and then check whether we end up with a consistent solution*. For the circuit in Fig. 4.6(a), we shall assume that both diodes are conducting. It follows that $V_B = 0$ and $V = 0$. The current through D_2 can now be determined from

$$I_{D2} = \frac{10-0}{10} = 1 \text{ mA}$$

Writing a node equation at B,

$$I + 1 = \frac{0-(-10)}{5}$$

results in $I = 1$ mA. Thus D_1 is conducting as originally assumed, and the final result is $I = 1$ mA and $V = 0$ V.

For the circuit in Fig. 4.6(b), if we assume that both diodes are conducting, then $V_B = 0$ and $V = 0$. The current in D_2 is obtained from

$$I_{D2} = \frac{10-0}{5} = 2 \text{ mA}$$

The node equation at B is

$$I + 2 = \frac{0-(-10)}{10}$$

which yields $I = -1$ mA. Since this is not possible, our original assumption is *not* correct. We start again, assuming that D_1 is off and D_2 is on. The current I_{D2} is given by

$$I_{D2} = \frac{10-(-10)}{15} = 1.33 \text{ mA}$$

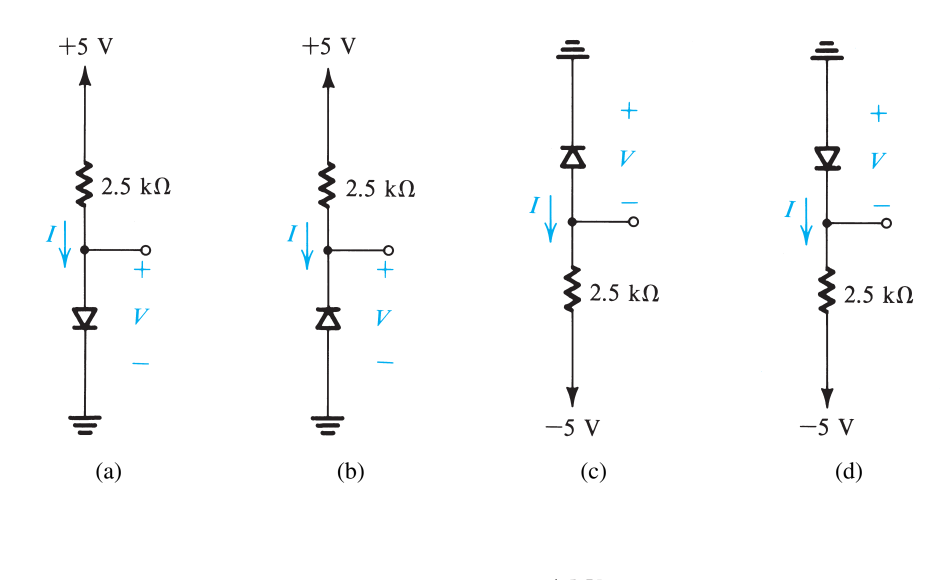
and the voltage at node B is

$$V_B = -10 + 10 \times 1.33 = +3.3 \text{ V}$$

Thus D_1 is reverse biased as assumed, and the final result is $I = 0$ and $V = 3.3$ V.

EXERCISES

4.4 Find the values of I and V in the circuits shown in Fig. E4.4.



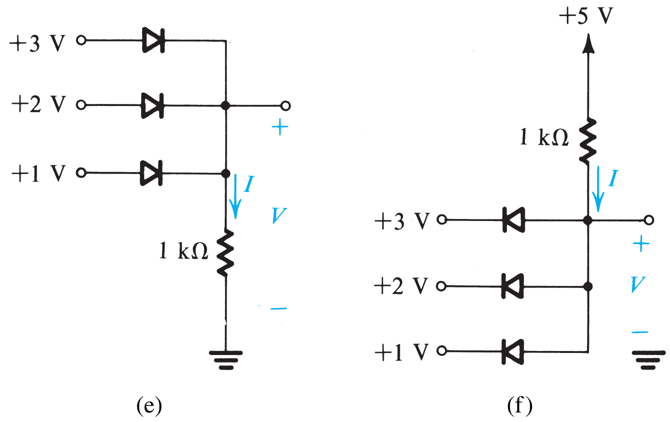


Figure E4.4

Ans. (a) 2 mA, 0 V; (b) 0 mA, 5 V; (c) 0 mA, 5 V; (d) 2 mA, 0 V; (e) 3 mA, +3 V; (f) 4 mA, +1 V

4.5 Figure E4.5 shows a circuit for an ac voltmeter. It utilizes a moving-coil meter that gives a full-scale reading when the *average* current flowing through it is 1 mA. The moving-coil meter has a 50-Ω resistance.

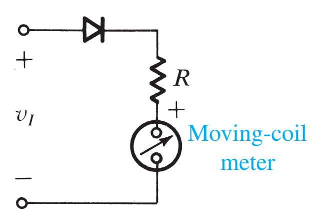


Figure E4.5

Find the value of R that results in the meter indicating a full-scale reading when the input sine-wave voltage v_I is 20 V peak-to-peak. (*Hint:* The average value of half-sine waves is V_p/π.)

Ans. 3.133 kΩ

4.2 Terminal Characteristics of Junction Diodes

The most common implementation of the diode utilizes a *pn* junction. We have studied the physics of the *pn* junction and derived its *i–v* characteristic in Chapter 3. That the *pn* junction is used to implement the diode function should come as no surprise: the *pn* junction can conduct substantial current in the forward direction and almost no current in the reverse direction. In this section we study the *i–v* characteristic of the *pn* junction diode in detail in order to prepare ourselves for diode circuit applications.

Figure 4.7 shows the *i–v* characteristic of a silicon junction diode. The same characteristic is shown in Fig. 4.8 with some scales expanded and others compressed to reveal details. Note that the scale changes have resulted in the apparent discontinuity at the origin.

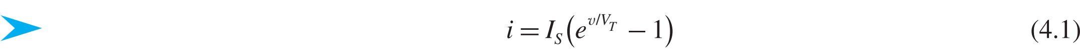
As indicated, the characteristic curve consists of three distinct regions:

1. The forward-bias region, determined by $v > 0$
2. The reverse-bias region, determined by $v < 0$
3. The breakdown region, determined by $v < -V_{ZK}$

These three regions of operation are described in the following sections.

4.2.1 The Forward-Bias Region

The forward-bias—or simply forward—region of operation is entered when the terminal voltage v is positive. In the forward region the *i–v* relationship is closely approximated by



$$i = I_S\left(e^{v/V_T} - 1\right) \tag{4.1}$$

In this equation[1] I_S is a constant for a given diode at a given temperature. A formula for I_S in terms of the diode's physical parameters and temperature was given in Eq. (3.41). The current

[1]Equation (4.1), the diode equation, is sometimes written to include a constant n in the exponential,

$$i = I_S(e^{v/nV_T} - 1)$$

with n having a value between 1 and 2, depending on the material and the physical structure of the diode. Diodes using the standard integrated-circuit fabrication process exhibit $n = 1$ when operated under normal conditions. For simplicity, we shall use $n = 1$ throughout this book, unless otherwise specified.

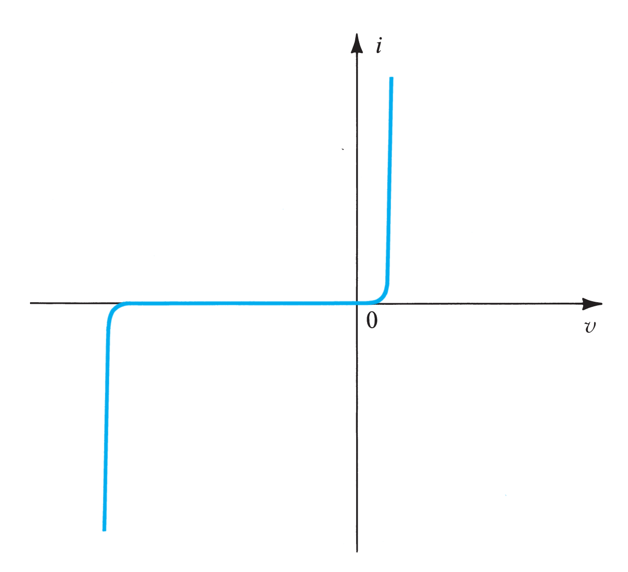


Figure 4.7 The i–v characteristic of a silicon junction diode.

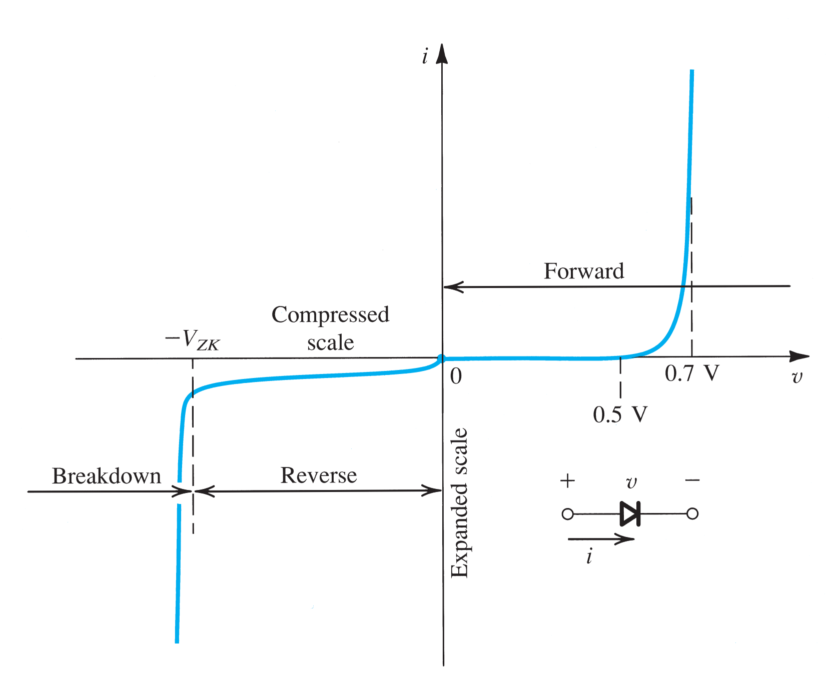


Figure 4.8 The diode i–v relationship with some scales expanded and others compressed in order to reveal details.

I_S is usually called the **saturation current** (for reasons that will become apparent shortly). Another name for I_S, and one that we will occasionally use, is the **scale current**. This name arises from the fact that I_S is directly proportional to the cross-sectional area of the diode. Thus doubling of the junction area results in a diode with double the value of I_S and, as the diode equation indicates, double the value of current i for a given forward voltage v. For "small-signal" diodes, which are small-size diodes intended for low-power applications, I_S is on the order of 10^{-15} A. The value of I_S is, however, a very strong function of temperature. As a rule of thumb, I_S doubles in value for every 5°C rise in temperature.

The voltage V_T in Eq. (4.1) is a constant called the **thermal voltage** and is given by

$$V_T = \frac{kT}{q} \tag{4.2}$$

where

k = Boltzmann's constant = 8.62×10^{-5} eV/K = 1.38×10^{-23} joules/kelvin
T = the absolute temperature in kelvins = 273 + temperature in °C
q = the magnitude of electronic charge = 1.60×10^{-19} coulomb

Substituting $k = 8.62 \times 10^{-5}$ eV/K into Eq. (4.2) gives

$$V_T = 0.0862T, \text{ mV} \tag{4.2a}$$

Thus, at room temperature (20°C) the value of V_T is 25.3 mV. In rapid approximate circuit analysis we shall use $V_T \simeq 25$ mV at room temperature.[2]

For appreciable current i in the forward direction, specifically for $i \gg I_S$, Eq. (4.1) can be approximated by the exponential relationship

$$i \simeq I_S e^{v/V_T} \tag{4.3}$$

This relationship can be expressed alternatively in the logarithmic form

$$v = V_T \ln \frac{i}{I_S} \tag{4.4}$$

where ln denotes the natural (base e) logarithm.

The exponential relationship of the current i to the voltage v holds over many decades of current (a span of as many as seven decades—i.e., a factor of 10^7—can be found). This is quite a remarkable property of junction diodes, one that is also found in bipolar junction transistors and that has been exploited in many interesting applications.

Let us consider the forward i–v relationship in Eq. (4.3) and evaluate the current I_1 corresponding to a diode voltage V_1:

$$I_1 = I_S e^{V_1/V_T}$$

Similarly, if the voltage is V_2, the diode current I_2 will be

$$I_2 = I_S e^{V_2/V_T}$$

[2]A slightly higher ambient temperature (25°C or so) is usually assumed for electronic equipment operating inside a cabinet. At this temperature, $V_T \simeq 25.8$ mV. Nevertheless, for the sake of simplicity and to promote rapid circuit analysis, we shall use the more arithmetically convenient value of $V_T \simeq 25$ mV throughout this book.

These two equations can be combined to produce

$$\frac{I_2}{I_1} = e^{(V_2 - V_1)/V_T}$$

which can be rewritten as

$$V_2 - V_1 = V_T \ln \frac{I_2}{I_1}$$

or, in terms of base-10 logarithms,

$$V_2 - V_1 = 2.3\, V_T \log \frac{I_2}{I_1} \tag{4.5}$$

This equation simply states that for a decade (factor of 10) change in current, the diode voltage drop changes by $2.3V_T$, which is approximately 60 mV. This also suggests that the diode i–v relationship is most conveniently plotted on semilog paper. Using the vertical, linear axis for v and the horizontal, log axis for i, one obtains a straight line with a slope of 60 mV per decade of current.

A glance at the i–v characteristic in the forward region (Fig. 4.8) reveals that the current is negligibly small for v smaller than about 0.5 V. This value is usually referred to as the **cut-in voltage**. It should be emphasized, however, that this apparent threshold in the characteristic is simply a consequence of the exponential relationship. Another consequence of this relationship is the rapid increase of i. Thus, for a "fully conducting" diode, the voltage drop lies in a narrow range, approximately 0.6 V to 0.8 V. This gives rise to a simple "model" for the diode where it is assumed that a conducting diode has approximately a 0.7-V drop across it. Diodes with different current ratings (i.e., different areas and correspondingly different I_S) will exhibit the 0.7-V drop at different currents. For instance, a small-signal diode may be considered to have a 0.7-V drop at $i = 1$ mA, while a higher-power diode may have a 0.7-V drop at $i = 1$ A. We will study the topics of diode-circuit analysis and diode models in the next section.

Example 4.3

A silicon diode said to be a 1-mA device displays a forward voltage of 0.7 V at a current of 1 mA. Evaluate the junction scaling constant I_S. What scaling constants would apply for a 1-A diode of the same manufacture that conducts 1 A at 0.7 V?

Solution

Since

$$i = I_S e^{v/V_T}$$

then

$$I_S = i e^{-v/V_T}$$

Example 4.3 *continued*

For the 1-mA diode:

$$I_S = 10^{-3}e^{-700/25} = 6.9 \times 10^{-16}\ \text{A}$$

The diode conducting 1 A at 0.7 V corresponds to one-thousand 1-mA diodes in parallel with a total junction area 1000 times greater. Thus I_S is also 1000 times greater,

$$I_S = 6.9 \times 10^{-13}\ \text{A}$$

Since both I_S and V_T are functions of temperature, the forward i–v characteristic varies with temperature, as illustrated in Fig. 4.9. At a given constant diode current, the voltage drop across the diode decreases by approximately 2 mV for every 1°C increase in temperature. The change in diode voltage with temperature has been exploited in the design of electronic thermometers.

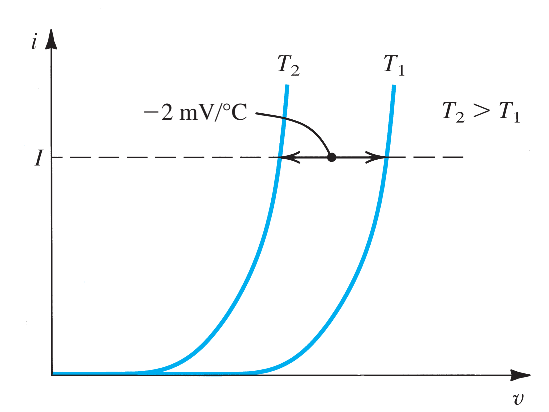


Figure 4.9 Temperature dependence of the diode forward characteristic. At a constant current, the voltage drop decreases by approximately 2 mV for every 1°C increase in temperature.

EXERCISES

4.6 Find the change in diode voltage if the current changes from 0.1 mA to 10 mA.
Ans. 120 mV

4.7 A silicon junction diode has $v = 0.7$ V at $i = 1$ mA. Find the voltage drop at $i = 0.1$ mA and $i = 10$ mA.
Ans. 0.64 V; 0.76 V

4.8 Using the fact that a silicon diode has $I_S = 10^{-14}$ A at 25°C and that I_S increases by 15% per °C rise in temperature, find the value of I_S at 125°C.
Ans. 1.17×10^{-8} A

4.2.2 The Reverse-Bias Region

The reverse-bias region of operation is entered when the diode voltage v is made negative. Equation (4.1) predicts that if v is negative and a few times larger than V_T (25 mV) in magnitude, the exponential term becomes negligibly small compared to unity, and the diode current becomes

$$i \simeq -I_S$$

That is, the current in the reverse direction is constant and equal to I_S. This constancy is the reason behind the term *saturation current*.

Real diodes exhibit reverse currents that, though quite small, are much larger than I_S. For instance, a small-signal diode whose I_S is on the order of 10^{-14} A to 10^{-15} A could show a reverse current on the order of 1 nA. The reverse current also increases somewhat with the increase in magnitude of the reverse voltage. Note that because of the very small magnitude of the current, these details are not clearly evident on the diode i–v characteristic of Fig. 4.8.

A large part of the reverse current is due to leakage effects. These leakage currents are proportional to the junction area, just as I_S is. Their dependence on temperature, however, is different from that of I_S. Thus, whereas I_S doubles for every 5°C rise in temperature, the corresponding rule of thumb for the temperature dependence of the reverse current is that it doubles for every 10°C rise in temperature.

EXERCISE

4.9 The diode in the circuit of Fig. E4.9 is a large high-current device whose reverse leakage is reasonably independent of voltage. If $V = 1$ V at 20°C, find the value of V at 40°C and at 0°C.

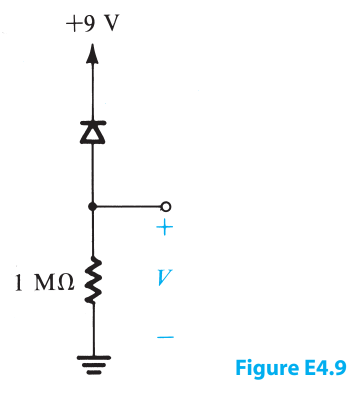


Figure E4.9

Ans. 4 V; 0.25 V

4.2.3 The Breakdown Region

The third distinct region of diode operation is the breakdown region, which can be easily identified on the diode i–v characteristic in Fig. 4.8. The breakdown region is entered when the magnitude of the reverse voltage exceeds a threshold value that is specific to the particular diode, called the **breakdown voltage**. This is the voltage at the "knee" of the i–v curve in Fig. 4.8 and is denoted V_{ZK}, where the subscript Z stands for zener (see Section 3.5.3) and K denotes knee.

As can be seen from Fig. 4.8, in the breakdown region the reverse current increases rapidly, with the associated increase in voltage drop being very small. Diode breakdown is normally not destructive, provided the power dissipated in the diode is limited by external circuitry to a "safe" level. This safe value is normally specified on the device data sheets. It therefore is necessary to limit the reverse current in the breakdown region to a value consistent with the permissible power dissipation.

The fact that the diode i–v characteristic in breakdown is almost a vertical line enables it to be used in voltage regulation. This subject will be studied in Section 4.5.

4.3 Modeling the Diode Forward Characteristic

Having studied the diode terminal characteristics we are now ready to consider the analysis of circuits employing forward-conducting diodes. Figure 4.10 shows such a circuit. It consists of a dc source V_{DD}, a resistor R, and a diode. We wish to analyze this circuit to determine the diode voltage V_D and current I_D. To aid in our analysis, we need to represent the diode with a model. There are a variety of diode models, of which we now know two: the ideal-diode model and the exponential model. In the following discussion we shall assess the suitability of these two models in various analysis situations. Also, we shall develop and comment on other models. This material, besides being useful in the analysis and design of diode circuits, establishes a foundation for the modeling of transistor operation that we will study in the next three chapters.

4.3.1 The Exponential Model

The most accurate description of the diode operation in the forward region is provided by the exponential model. Unfortunately, however, its severely nonlinear nature makes this model the most difficult to use. To illustrate, let's analyze the circuit in Fig. 4.10 using the exponential diode model.

Assuming that V_{DD} is greater than 0.5 V or so, the diode current will be much greater than I_S, and we can represent the diode i–v characteristic by the exponential relationship,

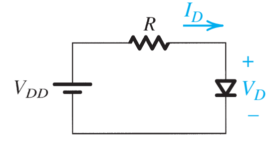


Figure 4.10 A simple circuit used to illustrate the analysis of circuits in which the diode is forward conducting.

resulting in

$$I_D = I_S e^{V_D/V_T} \tag{4.6}$$

The other equation that governs circuit operation is obtained by writing a Kirchhoff loop equation, resulting in

$$I_D = \frac{V_{DD} - V_D}{R} \tag{4.7}$$

Assuming that the diode parameter I_S is known, Eqs. (4.6) and (4.7) are two equations in the two unknown quantities I_D and V_D. Two alternative ways for obtaining the solution are graphical analysis and iterative analysis.

4.3.2 Graphical Analysis Using the Exponential Model

Graphical analysis is performed by plotting the relationships of Eqs. (4.6) and (4.7) on the i–v plane. The solution can then be obtained as the coordinates of the point of intersection of the two graphs. A sketch of the graphical construction is shown in Fig. 4.11. The curve represents the exponential diode equation (Eq. 4.6), and the straight line represents Eq. (4.7). Such a straight line is known as the **load line**, a name that will become more meaningful in later chapters. The load line intersects the diode curve at point Q, which represents the **operating point** of the circuit. Its coordinates give the values of I_D and V_D.

Graphical analysis aids in the visualization of circuit operation. However, the effort involved in performing such an analysis, particularly for complex circuits, is too great to be justified in practice.

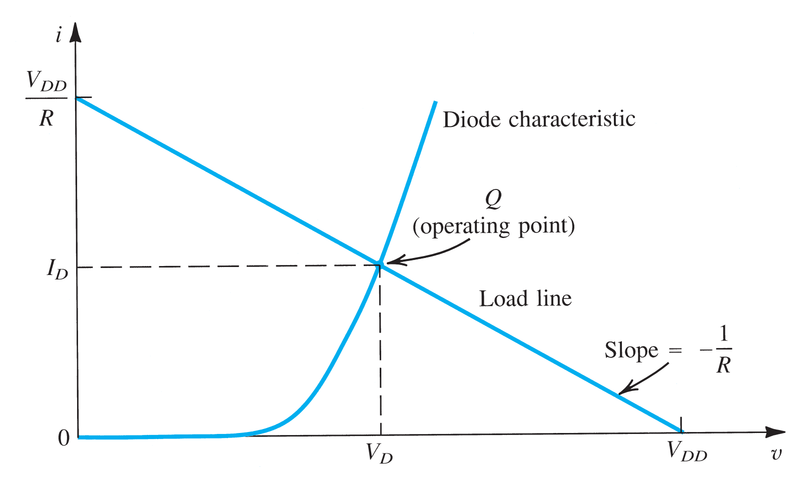


Figure 4.11 Graphical analysis of the circuit in Fig. 4.10 using the exponential diode model.

4.3.3 Iterative Analysis Using the Exponential Model

Equations (4.6) and (4.7) can be solved using a simple iterative procedure, as illustrated in the following example.

Example 4.4

Determine the current I_D and the diode voltage V_D for the circuit in Fig. 4.10 with $V_{DD} = 5$ V and $R = 1$ kΩ. Assume that the diode has a current of 1 mA at a voltage of 0.7 V.

Solution

To begin the iteration, we assume that $V_D = 0.7$ V and use Eq. (4.7) to determine the current,

$$I_D = \frac{V_{DD} - V_D}{R}$$

$$= \frac{5 - 0.7}{1} = 4.3 \text{ mA}$$

We then use the diode equation to obtain a better estimate for V_D. This can be done by employing Eq. (4.5), namely,

$$V_2 - V_1 = 2.3V_T \log \frac{I_2}{I_1}$$

Substituting $2.3V_T = 60$ mV, we have

$$V_2 = V_1 + 0.06 \log \frac{I_2}{I_1}$$

Substituting $V_1 = 0.7$ V, $I_1 = 1$ mA, and $I_2 = 4.3$ mA results in $V_2 = 0.738$ V. Thus the results of the first iteration are $I_D = 4.3$ mA and $V_D = 0.738$ V. The second iteration proceeds in a similar manner:

$$I_D = \frac{5 - 0.738}{1} = 4.262 \text{ mA}$$

$$V_2 = 0.738 + 0.06 \log \left[\frac{4.262}{4.3}\right]$$

$$= 0.738 \text{ V}$$

Thus the second iteration yields $I_D = 4.262$ mA and $V_D = 0.738$ V. Since these values are very close to the values obtained after the first iteration, no further iterations are necessary, and the solution is $I_D = 4.262$ mA and $V_D = 0.738$ V.

4.3.4 The Need for Rapid Analysis

The iterative analysis procedure utilized in the example above is simple and yields accurate results after two or three iterations. Nevertheless, there are situations in which the effort and time required are still greater than can be justified. Specifically, if one is doing a pencil-and-paper design of a relatively complex circuit, rapid circuit analysis is a necessity.

Through quick analysis, the designer is able to evaluate various possibilities before deciding on a suitable circuit design. To speed up the analysis process, one must be content with less precise results. This, however, is seldom a problem, because the more accurate analysis can be postponed until a final or almost-final design is obtained. Accurate analysis of the almost-final design can be performed with the aid of a computer circuit-analysis program such as SPICE (see Appendix B and the website). The results of such an analysis can then be used to further refine or "fine-tune" the design.

To speed up the analysis process, we must find a simpler model for the diode forward characteristic.

4.3.5 The Constant-Voltage-Drop Model

The simplest and most widely used diode model is the constant-voltage-drop model. This model is based on the observation that a forward-conducting diode has a voltage drop that varies in a relatively narrow range, say, 0.6 to 0.8 V. The model assumes this voltage to be constant at a value, say, 0.7 V. This development is illustrated in Fig. 4.12.

The constant-voltage-drop model is the one most frequently employed in the initial phases of analysis and design. This is especially true if at these stages one does not have detailed information about the diode characteristics, which is often the case.

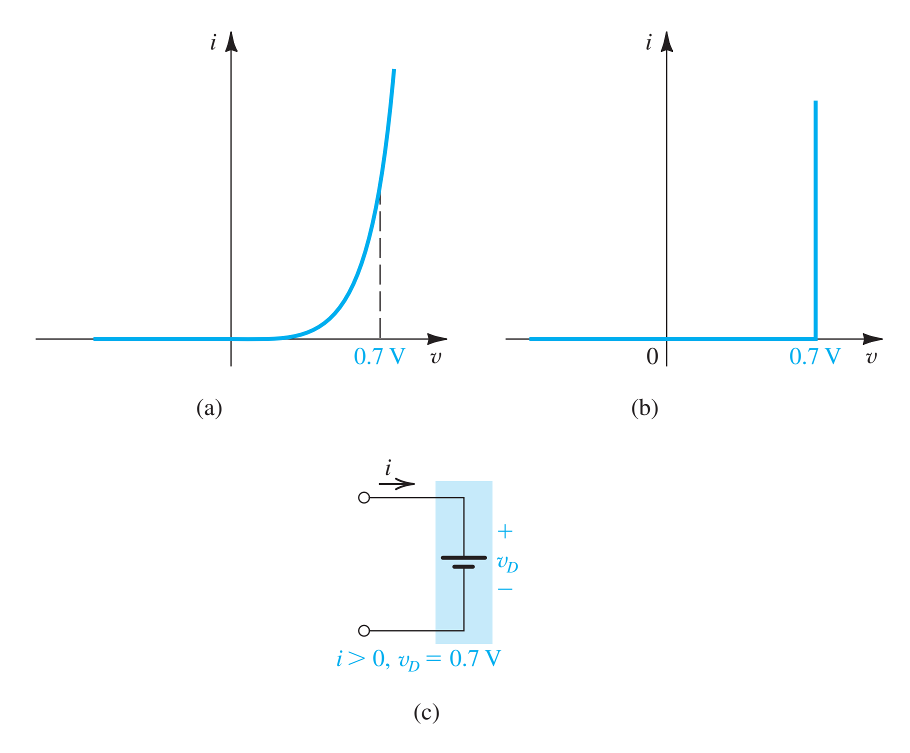


Figure 4.12 Development of the diode constant-voltage-drop model: **(a)** the exponential characteristic; **(b)** approximating the exponential characteristic by a constant voltage, usually about 0.7 V_i; **(c)** the resulting model of the forward-conducting diodes.

Finally, note that if we employ the constant-voltage-drop model to solve the problem in Example 4.4, we obtain

$$V_D = 0.7 \text{ V}$$

and

$$I_D = \frac{V_{DD} - 0.7}{R}$$
$$= \frac{5 - 0.7}{1} = 4.3 \text{ mA}$$

which are not very different from the values obtained before with the more elaborate exponential model.

4.3.6 The Ideal-Diode Model

In applications that involve voltages much greater than the diode voltage drop (0.6 V–0.8 V), we may neglect the diode voltage drop altogether while calculating the diode current. The result is the ideal-diode model, which we studied in Section 4.1. For the circuit in Example 4.4 (i.e., Fig. 4.10 with $V_{DD} = 5$ V and $R = 1$ kΩ), utilization of the ideal-diode model leads to

$$V_D = 0 \text{ V}$$
$$I_D = \frac{5 - 0}{1} = 5 \text{ mA}$$

which for a very quick analysis would not be bad as a gross estimate. However, with almost no additional work, the 0.7-V-drop model yields much more realistic results. We note, however, that the greatest utility of the ideal-diode model is in determining which diodes are on and which are off in a multidiode circuit, such as those considered in Section 4.1.

EXERCISES

4.10 For the circuit in Fig. 4.10, find I_D and V_D for the case $V_{DD} = 5$ V and $R = 10$ kΩ. Assume that the diode has a voltage of 0.7 V at 1-mA current. Use (a) iteration and (b) the constant-voltage-drop model with $V_D = 0.7$ V.
Ans. (a) 0.43 mA, 0.68 V; (b) 0.43 mA, 0.7 V

D4.11 Design the circuit in Fig. E4.11 to provide an output voltage of 2.4 V. Assume that the diodes available have 0.7-V drop at 1 mA.

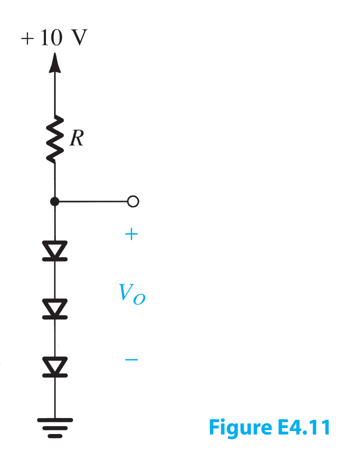


Figure E4.11

Ans. $R = 139\ \Omega$

4.12 Repeat Exercise 4.4 using the 0.7-V-drop model to obtain better estimates of I and V than those found in Exercise 4.4 (using the ideal-diode model).
Ans. (a) 1.72 mA, 0.7 V; (b) 0 mA, 5 V; (c) 0 mA, 5 V; (d) 1.72 mA, 0.7 V; (e) 2.3 mA, +2.3 V; (f) 3.3 mA, +1.7 V

4.3.7 The Small-Signal Model

Consider the situation in Fig. 4.13(a), where a dc voltage V_{DD} establishes a dc current I_D through the series combination of a resistance R and a diode D. The resulting diode voltage is denoted V_D. As mentioned above, values of I_D and V_D can be obtained by solving the circuit using the diode exponential characteristic or, much more quickly, approximate values can be found using the diode constant-voltage-drop model.

Next, consider the situation of V_{DD} undergoing a small change ΔV_{DD}, as shown in Fig. 4.13(b). As indicated, the current I_D changes by an increment ΔI_D, and the diode voltage V_D changes by an increment ΔV_D. We wish to find a quick way to determine the values of these incremental changes. Toward that end, we develop a "small-signal" model for the diode.

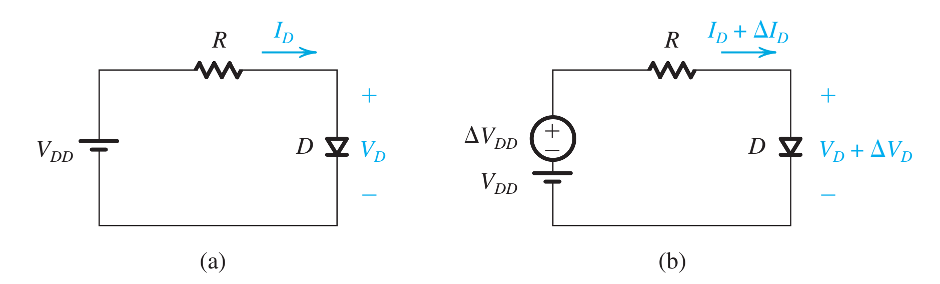


Figure 4.13 **(a)** A simple diode circuit; **(b)** the situation when V_{DD} changes by ΔV_{DD}.

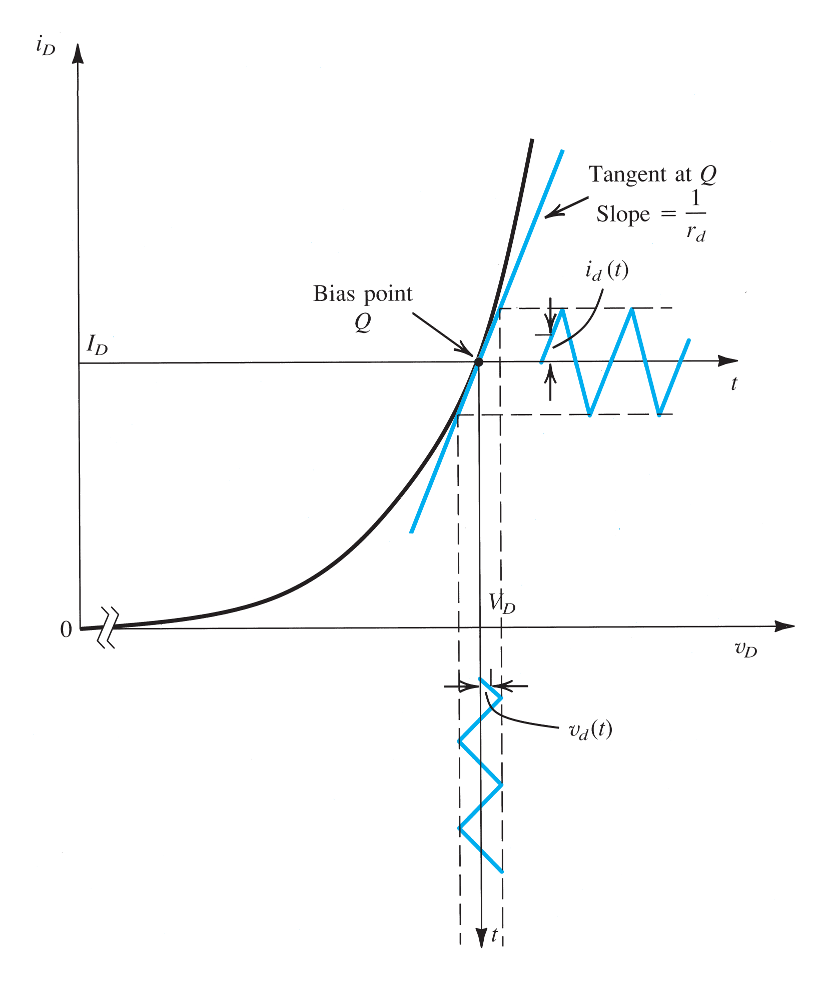


Figure 4.14 Development of the diode small-signal model.

Here the word *signal* emphasizes that in general, ΔV_{DD} can be a time-varying quantity. The qualifier "small" indicates that this diode model applies only when ΔV_D is kept sufficiently small, with "sufficiently" to be quantified shortly.

To develop the diode small-signal model, refer to Fig. 4.14. We express the voltage across the diode as the sum of the dc voltage V_D and the time-varying signal $v_d(t)$,

$$v_D(t) = V_D + v_d(t) \tag{4.8}$$

Correspondingly, the total instantaneous diode current $i_D(t)$ will be

$$i_D(t) = I_S e^{v_D/V_T} \tag{4.9}$$

Substituting for v_D from Eq. (4.8) gives

$$i_D(t) = I_S e^{(V_D + v_d)/V_T}$$

which can be rewritten

$$i_D(t) = I_S e^{V_D/V_T} e^{v_d/V_T} \tag{4.10}$$

In the absence of the signal $v_d(t)$, the diode voltage is equal to V_D, and the diode current is I_D, given by

$$I_D = I_S e^{V_D/V_T} \tag{4.11}$$

Thus, $i_D(t)$ in Eq. (4.10) can be expressed as

$$i_D(t) = I_D e^{v_d/V_T} \tag{4.12}$$

Now if the amplitude of the signal $v_d(t)$ is kept sufficiently small such that

$$\frac{v_d}{V_T} \ll 1 \tag{4.13}$$

then we may expand the exponential of Eq. (4.12) in a series and truncate the series after the first two terms to obtain the approximate expression

$$i_D(t) \simeq I_D\left(1 + \frac{v_d}{V_T}\right) \tag{4.14}$$

This is the **small-signal approximation**. It is valid for signals whose amplitudes are smaller than about 5 mV (see Eq. 4.13, and recall that $V_T = 25$ mV).[3]

From Eq. (4.14) we have

$$i_D(t) = I_D + \frac{I_D}{V_T} v_d \tag{4.15}$$

Thus, superimposed on the dc current I_D, we have a signal current component directly proportional to the signal voltage v_d. That is,

$$i_D = I_D + i_d \tag{4.16}$$

where

$$i_d = \frac{I_D}{V_T} v_d \tag{4.17}$$

The quantity relating the signal current i_d to the signal voltage v_d has the dimensions of conductance, mhos (℧), and is called the **diode small-signal conductance**. The inverse of this parameter is the **diode small-signal resistance**, or **incremental resistance**, r_d,

$$r_d = \frac{V_T}{I_D} \tag{4.18}$$

Note that the value of r_d is inversely proportional to the bias current I_D.

[3]For $v_d = 5$ mV, $v_d/V_T = 0.2$. Thus the next term in the series expansion of the exponential will be $\frac{1}{2} \times 0.2^2 = 0.02$, a factor of 10 lower than the linear term we kept.

Additional insight into the small-signal approximation and the small-signal diode model can be obtained by considering again the graphical construction in Fig. 4.14. Here the diode is seen to be operating at a dc bias point Q characterized by the dc voltage V_D and the corresponding dc current I_D. Superimposed on V_D we have a signal $v_d(t)$, assumed (arbitrarily) to have a triangular waveform.

It is easy to see that using the small-signal approximation is equivalent to assuming that *the signal amplitude is sufficiently small such that the excursion along the i–v curve is limited to a short almost-linear segment.* The slope of this segment, which is equal to the slope of the tangent to the i–v curve at the operating point Q, is equal to the small-signal conductance. The reader is encouraged to prove that the slope of the i–v curve at $i = I_D$ is equal to I_D/V_T, which is $1/r_d$; that is,

$$r_d = 1\bigg/\left[\frac{\partial i_D}{\partial v_D}\right]_{i_D=I_D} \qquad (4.19)$$

From the preceding we conclude that superimposed on the quantities V_D and I_D that define the dc bias point, or **quiescent point**, of the diode will be the small-signal quantities $v_d(t)$ and $i_d(t)$, which are related by the diode small-signal resistance r_d evaluated at the bias point (Eq. 4.18). Thus the small-signal analysis can be performed separately from the dc bias analysis, a great convenience that results from the linearization of the diode characteristics inherent in the small-signal approximation. Specifically, after the dc analysis is performed, the small-signal equivalent circuit is obtained by eliminating all dc sources (i.e., short-circuiting dc voltage sources and open-circuiting dc current sources) and replacing the diode by its small-signal resistance. Thus, for the circuit in Fig. 4.13(b), the dc analysis is obtained by using the circuit in Fig. 4.13(a), while the incremental quantities ΔI_D and ΔV_D can be determined by using the small-signal equivalent circuit shown in Fig. 4.15. The following example should further illustrate the application of the small-signal model.

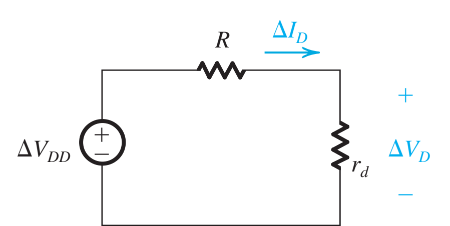


Figure 4.15 Circuit for determining the incremental quantities ΔI_D and ΔV_D for the circuit in Figure 4.13(b). Note that replacing the diode with its small-signal resistance r_d results in a linear circuit.

Example 4.5

Consider the circuit shown in Fig. 4.16(a) for the case in which $R = 10\ \text{k}\Omega$. The power supply V^+ has a dc value of 10 V on which is superimposed a 60-Hz sinusoid of 1-V peak amplitude. (This “signal” component of the power-supply voltage is an imperfection in the power-supply design. It is known as the **power-supply ripple**. More on this later.) Calculate both the dc voltage of the diode and the amplitude of the sine-wave signal appearing across it. Assume the diode to have a 0.7-V drop at 1-mA current.

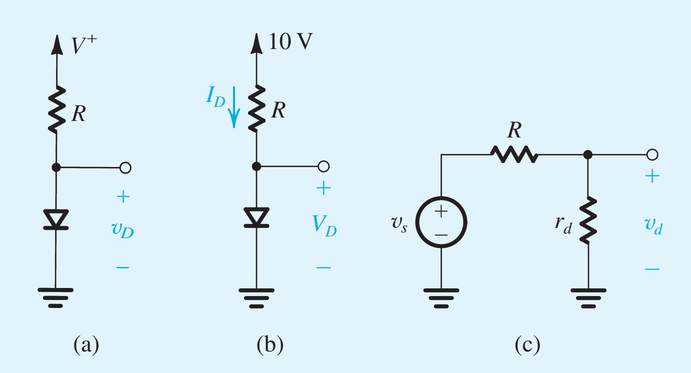


Figure 4.16 **(a)** Circuit for Example 4.5. **(b)** Circuit for calculating the dc operating point. **(c)** Small-signal equivalent circuit.

Solution

Considering dc quantities only, we assume $V_D \simeq 0.7$ V and calculate the diode dc current

$$I_D = \frac{10 - 0.7}{10} = 0.93 \text{ mA}$$

Since this value is very close to 1 mA, the diode voltage will be very close to the assumed value of 0.7 V. At this operating point, the diode incremental resistance r_d is

$$r_d = \frac{V_T}{I_D} = \frac{25}{0.93} = 26.9\ \Omega$$

The signal voltage across the diode can be found from the small-signal equivalent circuit in Fig. 4.16(c). Here v_s denotes the 60-Hz 1-V peak sinusoidal component of V^+, and v_d is the corresponding signal across the diode. Using the voltage divider rule provides the peak amplitude of v_d as follows:

$$v_d(\text{peak}) = \hat{V}_s \frac{r_d}{R + r_d}$$

$$= 1 \frac{0.0269}{10 + 0.0269} = 2.68 \text{ mV}$$

Finally, we note that since this value is quite small, our use of the small-signal model of the diode is justified.

From the above we see that for a diode circuit that involves both dc and signal quantities, a small-signal equivalent circuit can be obtained by eliminating the dc sources and replacing each diode with its small-signal resistance r_d. Such a circuit is linear and can be solved using linear circuit analysis.

Finally, we note that while r_d models the small-signal operation of the diode at low frequencies, its dynamic operation is modeled by the capacitances C_j and C_d, which we

studied in Section 3.6 and which also are small-signal parameters. A complete model of the diode includes C_j and C_d in parallel with r_d.

4.3.8 Use of the Diode Forward Drop in Voltage Regulation

A further application of the diode small-signal model is found in a popular diode application, namely, the use of diodes to create a regulated voltage. A **voltage regulator** is a circuit whose purpose is to provide a constant dc voltage between its output terminals. The output voltage is required to remain as constant as possible in spite of (a) changes in the load current drawn from the regulator output terminal and (b) changes in the dc power-supply voltage that feeds the regulator circuit. Since the forward-voltage drop of the diode remains almost constant at approximately 0.7 V while the current through it varies by relatively large amounts, a forward-biased diode can make a simple voltage regulator. For instance, we have seen in Example 4.5 that while the 10-V dc supply voltage had a ripple of 2 V peak-to-peak (a ±10% variation), the corresponding ripple in the diode voltage was only about ±2.7 mV (a ±0.4% variation). Regulated voltages greater than 0.7 V can be obtained by connecting a number of diodes in series. For example, the use of three forward-biased diodes in series provides a voltage of about 2 V. One such circuit is investigated in the following example, which utilizes the diode small-signal model to quantify the efficacy of the voltage regulator that is realized.

Example 4.6

Consider the circuit shown in Fig. 4.17. A string of three diodes is used to provide a constant voltage of about 2.1 V. We want to calculate the percentage change in this regulated voltage caused by (a) a ±10% change in the power-supply voltage, and (b) connection of a 1-kΩ load resistance.

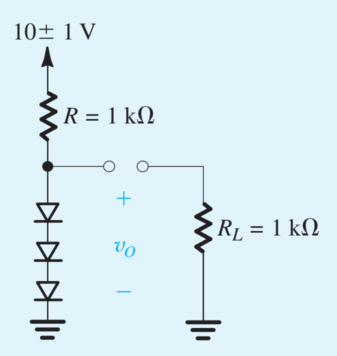


Figure 4.17 Circuit for Example 4.6.

Solution

With no load, the nominal value of the current in the diode string is given by

$$I = \frac{10 - 2.1}{1} = 7.9 \text{ mA}$$

Thus each diode will have an incremental resistance of

$$r_d = \frac{V_T}{I}$$

Thus,

$$r_d = \frac{25}{7.9} = 3.2\ \Omega$$

The three diodes in series will have a total incremental resistance of

$$r = 3r_d = 9.6\ \Omega$$

This resistance, along with the resistance R, forms a voltage divider whose ratio can be used to calculate the change in output voltage due to a $\pm10\%$ (i.e., ±1-V) change in supply voltage. Thus the peak-to-peak change in output voltage will be

$$\Delta v_O = 2\frac{r}{r+R} = 2\frac{0.0096}{0.0096+1} = 19 \text{ mV peak-to-peak}$$

That is, corresponding to the ±1-V ($\pm10\%$) change in supply voltage, the output voltage will change by ±9.5 mV or $\pm0.5\%$. Since this implies a change of about ±3.2 mV per diode, our use of the small-signal model is justified.

When a load resistance of 1 kΩ is connected across the diode string, it draws a current of approximately 2.1 mA. Thus the current in the diodes decreases by 2.1 mA, resulting in a decrease in voltage across the diode string given by

$$\Delta v_O = -2.1 \times r = -2.1 \times 9.6 = -20 \text{ mV}$$

Since this implies that the voltage across each diode decreases by about 6.7 mV, our use of the small-signal model is not entirely justified. Nevertheless, a detailed calculation of the voltage change using the exponential model results in $\Delta v_O = -23$ mV, which is not too different from the approximate value obtained using the incremental model.

EXERCISES

4.13 Find the value of the diode small-signal resistance r_d at bias currents of 0.1 mA, 1 mA, and 10 mA.
Ans. 250 Ω; 25 Ω; 2.5 Ω

4.14 Consider a diode biased at 1 mA. Find the change in current as a result of changing the voltage by (a) –10 mV, (b) –5 mV, (c) +5 mV, and (d) +10 mV. In each case, do the calculations (i) using the small-signal model and (ii) using the exponential model.
Ans. (a) –0.40, –0.33 mA; (b) –0.20, –0.18 mA; (c) +0.20, +0.22 mA; (d) +0.40, +0.49 mA

D4.15 Design the circuit of Fig. E4.15 so that $V_O = 3$ V when $I_L = 0$, and V_O changes by 20 mV per 1 mA of load current.
(a) Use the small-signal model of the diode to find the value of R.
(b) Specify the value of I_S of each of the diodes.

(c) For this design, use the diode exponential model to determine the actual change in V_O when a current $I_L = 1$ mA is drawn from the regulator.

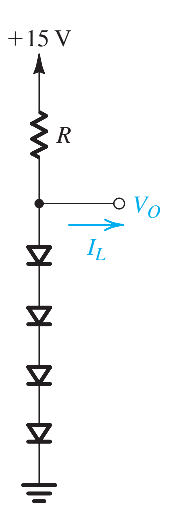


Figure E4.15

Ans. (a) $R = 2.4\ \text{k}\Omega$; (b) $I_S = 4.7 \times 10^{-16}$ A; (c) –23 mV

4.4 Operation in the Reverse Breakdown Region—Zener Diodes

The very steep i–v curve that the diode exhibits in the breakdown region (Fig. 4.8) and the almost-constant voltage drop that this indicates suggest that diodes operating in the breakdown region can be used in the design of voltage regulators. From the previous section, the reader will recall that voltage regulators are circuits that provide a constant dc output voltage in the face of changes in their load current and in the system power-supply voltage. This in fact turns out to be an important application of diodes operating in the reverse breakdown region, and special diodes are manufactured to operate specifically in the breakdown region. Such diodes are called **breakdown diodes** or, more commonly, as noted earlier, **zener diodes**.

Figure 4.18 shows the circuit symbol of the zener diode. In normal applications of zener diodes, current flows into the cathode, and the cathode is positive with respect to the anode. Thus I_Z and V_Z in Fig. 4.18 have positive values.

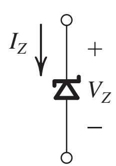


Figure 4.18 Circuit symbol for a zener diode.

4.4.1 Specifying and Modeling the Zener Diode

Figure 4.19 shows details of the diode i–v characteristic in the breakdown region. We observe that for currents greater than the **knee current** I_{ZK} (specified on the data sheet of the zener diode), the i–v characteristic is almost a straight line. The manufacturer usually specifies the voltage across the zener diode V_Z at a specified test current, I_{ZT}. We have indicated these parameters in Fig. 4.19 as the coordinates of the point labeled Q. Thus a 6.8-V zener diode will exhibit a 6.8-V drop at a specified test current of, say, 10 mA. As the current through the zener deviates from I_{ZT}, the voltage across it will change, though only slightly. Figure 4.19 shows that corresponding to current change ΔI the zener voltage changes by ΔV, which is related to ΔI by

$$\Delta V = r_z \Delta I$$

where r_z is the inverse of the slope of the almost-linear i–v curve at point Q. Resistance r_z is the **incremental resistance** of the zener diode at operating point Q. It is also known as the **dynamic resistance** of the zener, and its value is specified on the device data sheet. Typically, r_z is in the range of a few ohms to a few tens of ohms. Obviously, the lower the value of r_z is, the more constant the zener voltage remains as its current varies, and thus the more ideal its performance becomes in the design of voltage regulators. In this regard, we observe from Fig. 4.19 that while r_z remains low and almost constant over a wide range of current, its value increases considerably in the vicinity of the knee. Therefore, as a general design guideline, one should avoid operating the zener in this low-current region.

Zener diodes are fabricated with voltages V_Z in the range of a few volts to a few hundred volts. In addition to specifying V_Z (at a particular current I_{ZT}), r_z, and I_{ZK}, the manufacturer

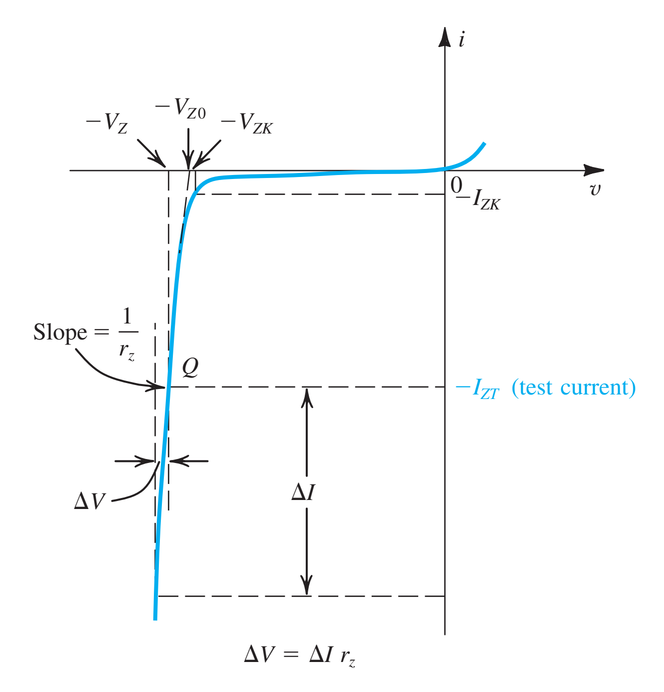


Figure 4.19 The diode i–v characteristic with the breakdown region shown in some detail.

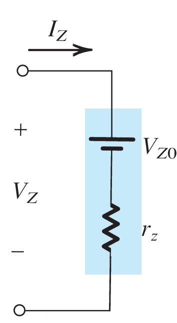


Figure 4.20 Model for the zener diode.

also specifies the maximum power that the device can safely dissipate. Thus a 0.5-W, 6.8-V zener diode can operate safely at currents up to a maximum of about 70 mA.

The almost-linear i–v characteristic of the zener diode suggests that the device can be modeled as indicated in Fig. 4.20. Here V_{Z0} denotes the point at which the straight line of slope $1/r_z$ intersects the voltage axis (refer to Fig. 4.19). Although V_{Z0} is shown in Fig. 4.19 to be slightly different from the knee voltage V_{ZK}, in practice their values are almost equal. The equivalent circuit model of Fig. 4.20 can be analytically described by

$$V_Z = V_{Z0} + r_z I_Z \tag{4.20}$$

and it applies for $I_Z > I_{ZK}$ and, obviously, $V_Z > V_{Z0}$.

4.4.2 Use of the Zener as a Shunt Regulator

We now illustrate, by way of an example, the use of zener diodes in the design of shunt regulators, so named because the regulator circuit appears in parallel (shunt) with the load.

Example 4.7

The 6.8-V zener diode in the circuit of Fig. 4.21(a) is specified to have $V_Z = 6.8$ V at $I_Z = 5$ mA, $r_z = 20\ \Omega$, and $I_{ZK} = 0.2$ mA. The supply voltage V^+ is nominally 10 V but can vary by ±1 V.

(a) Find V_O with no load and with V^+ at its nominal value.
(b) Find the change in V_O resulting from the ±1-V change in V^+. Note that $\left(\Delta V_O/\Delta V^+\right)$, usually expressed in mV/V, is known as **line regulation**.
(c) Find the change in V_O resulting from connecting a load resistance R_L that draws a current $I_L = 1$ mA, and hence find the **load regulation** $\left(\Delta V_O/\Delta I_L\right)$ in mV/mA.
(d) Find the change in V_O when $R_L = 2$ kΩ.
(e) Find the value of V_O when $R_L = 0.5$ kΩ.
(f) What is the minimum value of R_L for which the diode still operates in the breakdown region?

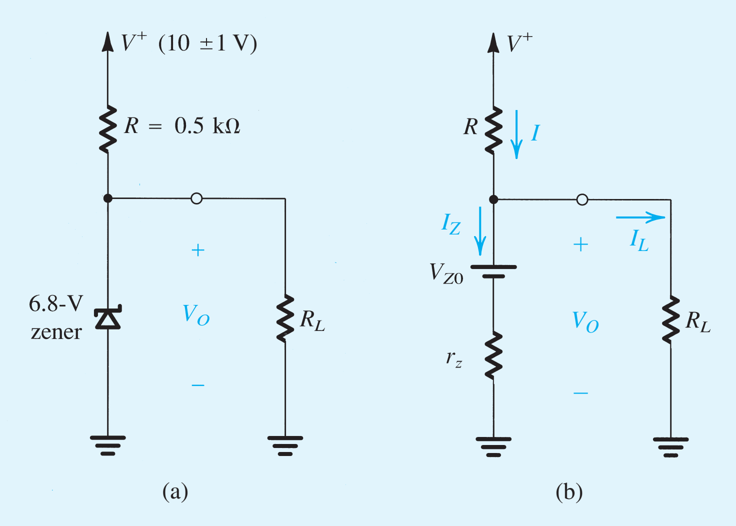


Figure 4.21 **(a)** Circuit for Example 4.7. **(b)** The circuit with the zener diode replaced with its equivalent circuit model.

Solution

First we must determine the value of the parameter V_{Z0} of the zener diode model. Substituting $V_Z = 6.8$ V, $I_Z = 5$ mA, and $r_z = 20\ \Omega$ in Eq. (4.20) yields $V_{Z0} = 6.7$ V. Figure 4.21(b) shows the circuit with the zener diode replaced with its model.

(a) With no load connected, the current through the zener is given by

$$I_Z = I = \frac{V^+ - V_{Z0}}{R + r_z}$$

$$= \frac{10 - 6.7}{0.5 + 0.02} = 6.35 \text{ mA}$$

Thus,

$$V_O = V_{Z0} + I_Z r_z$$

$$= 6.7 + 6.35 \times 0.02 = 6.83 \text{ V}$$

(b) For a ±1-V change in V^+, the change in output voltage can be found from

$$\Delta V_O = \Delta V^+ \frac{r_z}{R + r_z}$$

$$= \pm 1 \times \frac{20}{500 + 20} = \pm 38.5 \text{ mV}$$

Thus,

$$\text{Line regulation} = 38.5 \text{ mV/V}$$

Example 4.7 *continued*

(c) When a load resistance R_L that draws a load current $I_L = 1$ mA is connected, the zener current will decrease by 1 mA. The corresponding change in zener voltage can be found from

$$\Delta V_O = r_z \Delta I_Z$$
$$= 20 \times -1 = -20 \text{ mV}$$

Thus the load regulation is

$$\text{Load regulation} \equiv \frac{\Delta V_O}{\Delta I_L} = -20 \text{ mV/mA}$$

(d) When a load resistance of 2 kΩ is connected, the load current will be approximately 6.8 V/2 kΩ = 3.4 mA. Thus the change in zener current will be $\Delta I_Z = -3.4$ mA, and the corresponding change in zener voltage (output voltage) will thus be

$$\Delta V_O = r_z \Delta I_Z$$
$$= 20 \times -3.4 = -68 \text{ mV}$$

This value could have been obtained by multiplying the load regulation by the value of I_L (3.4 mA).

(e) An R_L of 0.5 kΩ would draw a load current of 6.8/0.5 = 13.6 mA. This is not possible, because the current I supplied through R is only 6.4 mA (for $V^+ = 10$ V). Therefore, the zener must be cut off. If this is indeed the case, then V_O is determined by the voltage divider formed by R_L and R (Fig. 4.21a),

$$V_O = V^+ \frac{R_L}{R + R_L}$$
$$= 10 \frac{0.5}{0.5 + 0.5} = 5 \text{ V}$$

Since this voltage is lower than the breakdown voltage of the zener, the diode is indeed no longer operating in the breakdown region.

(f) For the zener to be at the edge of the breakdown region, $I_Z = I_{ZK} = 0.2$ mA and $V_Z \simeq V_{ZK} \simeq 6.7$ V. At this point the lowest (worst-case) current supplied through R is $(9 - 6.7)/0.5 = 4.6$ mA, and thus the load current is $4.6 - 0.2 = 4.4$ mA. The corresponding value of R_L is

$$R_L = \frac{6.7}{4.4} \simeq 1.5 \text{ k}\Omega$$

4.4.3 Temperature Effects

The dependence of the zener voltage V_Z on temperature is specified in terms of its **temperature coefficient TC**, or **temco** as it is commonly known, which is usually expressed in mV/°C. The value of TC depends on the zener voltage, and for a given diode the TC varies with the operating current. Zener diodes whose V_Z are lower than about 5 V exhibit a negative TC. On the other hand, zeners with higher voltages exhibit a positive TC. The TC of a zener diode with a V_Z of about 5 V can be made zero by operating the diode at a specified current. Another commonly used technique for obtaining a reference voltage with low temperature coefficient

is to connect a zener diode with a positive temperature coefficient of about 2 mV/°C in series with a forward-conducting diode. Since the forward-conducting diode has a voltage drop of $\simeq$0.7 V and a TC of about –2 mV/°C, the series combination will provide a voltage of $(V_Z + 0.7)$ with a TC of about zero.

EXERCISES

4.16 A zener diode whose nominal voltage is 10 V at 10 mA has an incremental resistance of 50 Ω. What voltage do you expect if the diode current is halved? Doubled? What is the value of V_{Z0} in the zener model?
Ans. 9.75 V; 10.5 V; 9.5 V

4.17 A zener diode exhibits a constant voltage of 5.6 V for currents greater than five times the knee current. I_{ZK} is specified to be 1 mA. The zener is to be used in the design of a shunt regulator fed from a 15-V supply. The load current varies over the range of 0 mA to 15 mA. Find a suitable value for the resistor R. What is the maximum power dissipation of the zener diode?
Ans. 470 Ω; 112 mW

4.18 A shunt regulator utilizes a zener diode whose voltage is 5.1 V at a current of 50 mA and whose incremental resistance is 7 Ω. The diode is fed from a supply of 15-V nominal voltage through a 200-Ω resistor. What is the output voltage at no load? Find the line regulation and the load regulation.
Ans. 5.1 V; 33.8 mV/V; –7 mV/mA

4.4.4 A Final Remark

Though simple and useful, zener diodes have lost a great deal of their popularity in recent years. They have been virtually replaced in voltage-regulator design by specially designed integrated circuits (ICs) that perform the voltage-regulation function much more effectively and with greater flexibility than zener diodes.

4.5 Rectifier Circuits

One of the most important applications of diodes is in the design of rectifier circuits. A diode rectifier forms an essential building block of the dc power supplies required to power electronic equipment. A block diagram of such a power supply is shown in Fig. 4.22. As indicated, the power supply is fed from the 120-V (rms) 60-Hz ac line, and it delivers a dc voltage V_O (usually in the range of 4 V to 20 V) to an electronic circuit represented by the *load* block. The dc voltage V_O is required to be as constant as possible in spite of variations in the ac line voltage and in the current drawn by the load.

The first block in a dc power supply is the **power transformer**. It consists of two separate coils wound around an iron core that magnetically couples the two windings. The **primary winding**, having N_1 turns, is connected to the 120-V ac supply, and the **secondary winding**, having N_2 turns, is connected to the circuit of the dc power supply. Thus an ac voltage v_S of $120(N_2/N_1)$ V (rms) develops between the two terminals of the secondary winding. By

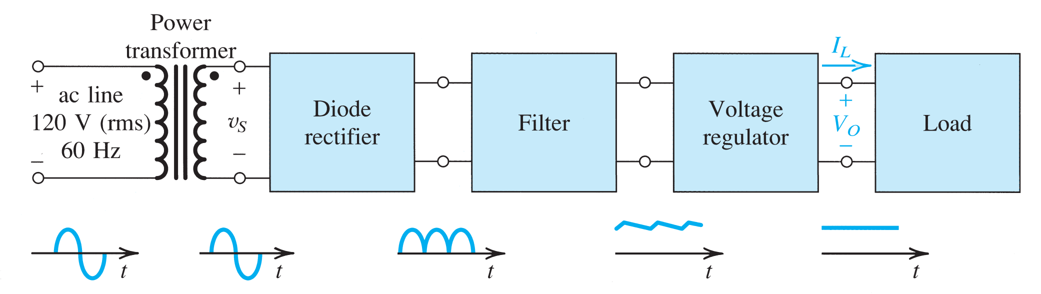


Figure 4.22 Block diagram of a dc power supply.

selecting an appropriate turns ratio (N_1/N_2) for the transformer, the designer can step the line voltage down to the value required to yield the particular dc voltage output of the supply. For instance, a secondary voltage of 8-V rms may be appropriate for a dc output of 5 V. This can be achieved with a 15:1 turns ratio.

In addition to providing the appropriate sinusoidal amplitude for the dc power supply, the power transformer provides electrical isolation between the electronic equipment and the power-line circuit. This isolation minimizes the risk of electric shock to the equipment user.

The diode rectifier converts the input sinusoid v_S to a unipolar output, which can have the pulsating waveform indicated in Fig. 4.22. Although this waveform has a nonzero average or a dc component, its pulsating nature makes it unsuitable as a dc source for electronic circuits, hence the need for a filter. The variations in the magnitude of the rectifier output are considerably reduced by the filter block in Fig. 4.22. In this section we shall study a number of rectifier circuits and a simple implementation of the output filter.

The output of the rectifier filter, though much more constant than without the filter, still contains a time-dependent component, known as **ripple**. To reduce the ripple and to stabilize the magnitude of the dc output voltage against variations caused by changes in load current, a voltage regulator is employed. Such a regulator can be implemented using the zener shunt regulator configuration studied in Section 4.4. Alternatively, and much more commonly at present, an integrated-circuit regulator can be used.

4.5.1 The Half-Wave Rectifier

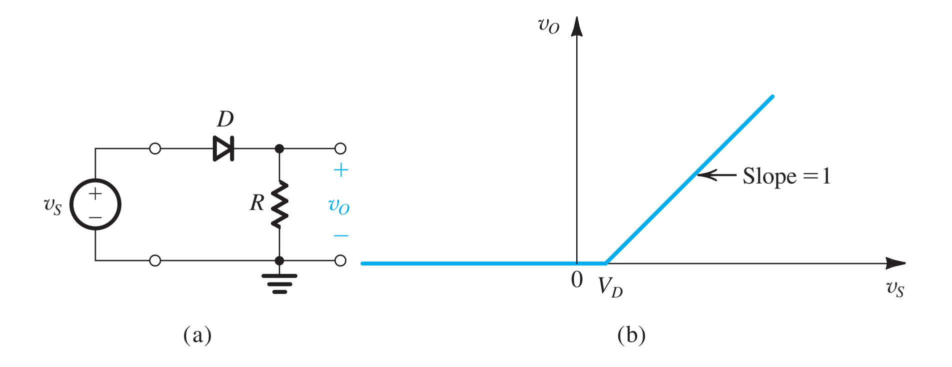
The half-wave rectifier utilizes alternate half-cycles of the input sinusoid. Figure 4.23(a) shows the circuit of a half-wave rectifier. This circuit was analyzed in Section 4.1 (see Fig. 4.3) assuming an ideal diode. Using the more realistic constant-voltage-drop diode model, we obtain

$$v_O = 0, \qquad v_S < V_D \tag{4.21a}$$

$$v_O = v_S - V_D, \qquad v_S \geq V_D \tag{4.21b}$$

The transfer characteristic represented by these equations is sketched in Fig. 4.23(b), where $V_D = 0.7$ V or 0.8 V. Figure 4.23(c) shows the output voltage obtained when the input v_S is a sinusoid.

In selecting diodes for rectifier design, two important parameters must be specified: the current-handling capability required of the diode, determined by the largest current the diode is expected to conduct, and the **peak inverse voltage** (PIV) that the diode must be able to



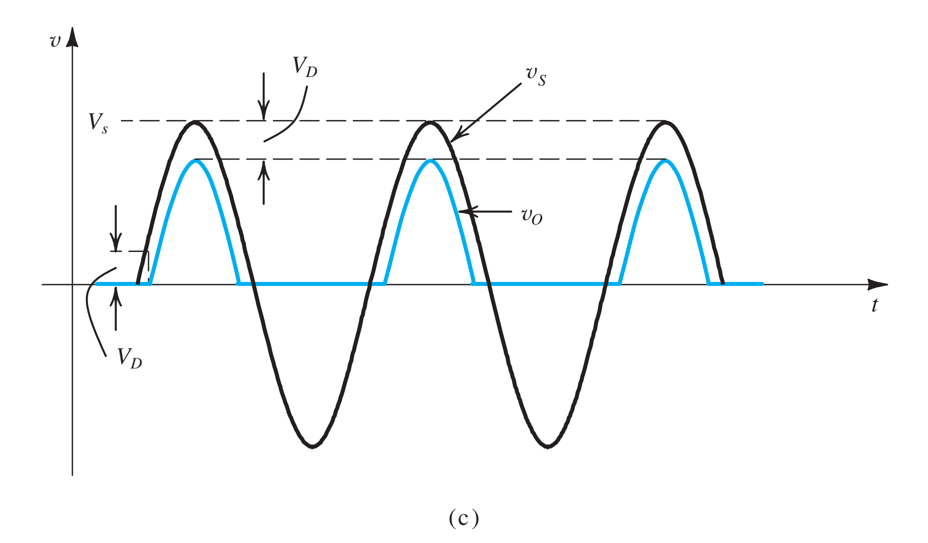


Figure 4.23 **(a)** Half-wave rectifier. **(b)** Transfer characteristic of the rectifier circuit. **(c)** Input and output waveforms.

withstand without breakdown, determined by the largest reverse voltage that is expected to appear across the diode. In the rectifier circuit of Fig. 4.23(a), we observe that when v_S is negative the diode will be cut off and v_O will be zero. It follows that the PIV is equal to the peak of v_S,

$$\text{PIV} = V_s \tag{4.22}$$

It is usually prudent, however, to select a diode that has a reverse breakdown voltage at least 50% greater than the expected PIV.

Before leaving the half-wave rectifier, the reader should note two points. First, it is possible to use the diode exponential characteristic to determine the exact transfer characteristic of the rectifier (see Problem 4.68). However, the amount of work involved is usually too great to be justified in practice. Of course, such an analysis can be easily done using a computer circuit-analysis program such as SPICE.

Second, whether we analyze the circuit accurately or not, it should be obvious that this circuit does not function properly when the input signal is small. For instance, this circuit cannot be used to rectify an input sinusoid of 100-mV amplitude. For such an application one

resorts to a so-called precision rectifier, a circuit utilizing diodes in conjunction with op amps. One such circuit is presented in Section 4.5.5.

EXERCISE

4.19 For the half-wave rectifier circuit in Fig. 4.23(a), show the following: (a) For the half-cycles during which the diode conducts, conduction begins at an angle $\theta = \sin^{-1}(V_D/V_s)$ and terminates at $(\pi - \theta)$, for a total conduction angle of $(\pi - 2\theta)$. (b) The average value (dc component) of v_O is $V_O \simeq (1/\pi)V_s - V_D/2$. (c) The peak diode current is $(V_s - V_D)/R$.
Find numerical values for these quantities for the case of 12-V (rms) sinusoidal input, $V_D \simeq 0.7$ V, and $R = 100\ \Omega$. Also, give the value for PIV.
Ans. (a) $\theta = 2.4°$, conduction angle $= 175°$; (b) 5.05 V; (c) 163 mA; 17 V

4.5.2 The Full-Wave Rectifier

The full-wave rectifier utilizes both halves of the input sinusoid. To provide a unipolar output, it inverts the negative halves of the sine wave. One possible implementation is shown in Fig. 4.24(a). Here the transformer secondary winding is **center-tapped** to provide two equal voltages v_S across the two halves of the secondary winding with the polarities indicated. Note that when the input line voltage (feeding the primary) is positive, both of the signals labeled v_S will be positive. In this case D_1 will conduct and D_2 will be reverse biased. The current through D_1 will flow through R and back to the center tap of the secondary. The circuit then behaves like a half-wave rectifier, and the output during the positive half-cycles when D_1 conducts will be identical to that produced by the half-wave rectifier.

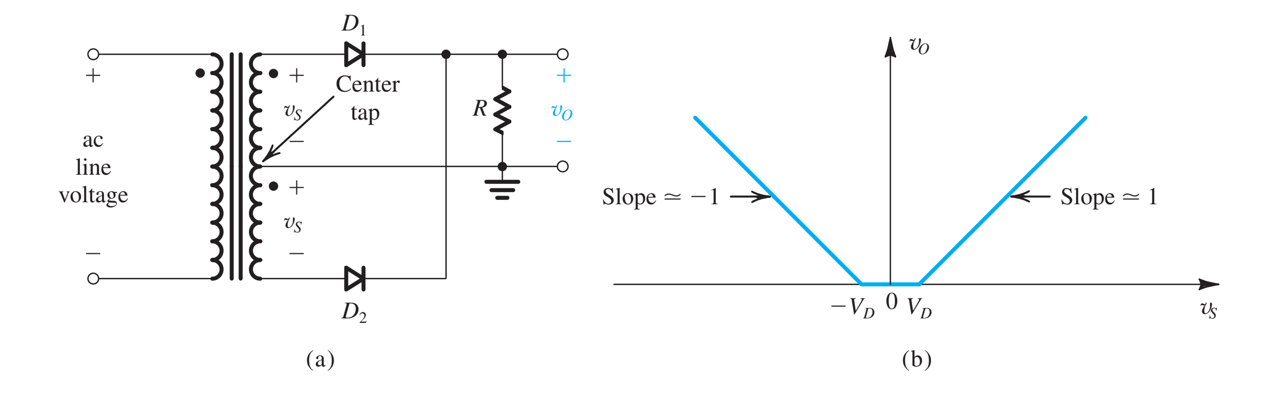
Now, during the negative half-cycle of the ac line voltage, both of the voltages labeled v_S will be negative. Thus D_1 will be cut off while D_2 will conduct. The current conducted by D_2 will flow through R and back to the center tap. It follows that during the negative half-cycles while D_2 conducts, the circuit behaves again as a half-wave rectifier. The important point, however, is that the current through R always flows in the same direction, and thus v_O will be unipolar, as indicated in Fig. 4.24(c). The output waveform shown is obtained by assuming that a conducting diode has a constant voltage drop V_D. Thus the transfer characteristic of the full-wave rectifier takes the shape shown in Fig. 4.24(b).

The full-wave rectifier obviously produces a more "energetic" waveform than that provided by the half-wave rectifier. In almost all rectifier applications, one opts for a full-wave type of some kind.

To find the PIV of the diodes in the full-wave rectifier circuit, consider the situation during the positive half-cycles. Diode D_1 is conducting, and D_2 is cut off. The voltage at the cathode of D_2 is v_O, and that at its anode is $-v_S$. Thus the reverse voltage across D_2 will be $(v_O + v_S)$, which will reach its maximum when v_O is at its peak value of $(V_s - V_D)$, and v_S is at its peak value of V_s; thus,

$$\text{PIV} = 2V_s - V_D$$

which is approximately twice that for the case of the half-wave rectifier.



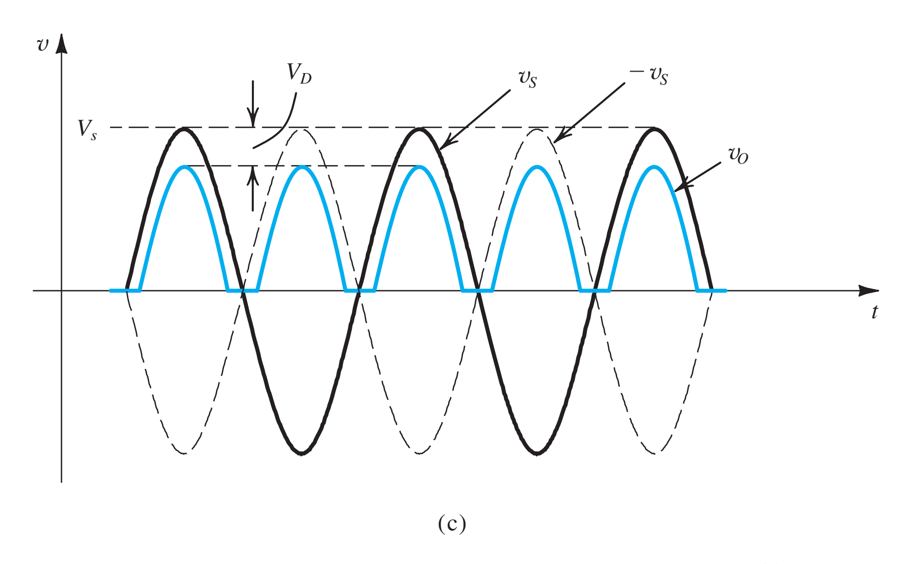


Figure 4.24 Full-wave rectifier utilizing a transformer with a center-tapped secondary winding: **(a)** circuit; **(b)** transfer characteristic assuming a constant-voltage-drop model for the diodes; **(c)** input and output waveforms.

EXERCISE

4.20 For the full-wave rectifier circuit in Fig. 4.24(a), show the following: (a) The output is zero for an angle of $2\sin^{-1}(V_D/V_s)$ centered around the zero-crossing points of the sine-wave input. (b) The average value (dc component) of v_O is $V_O \simeq (2/\pi)V_s - V_D$. (c) The peak current through each diode is $(V_s - V_D)/R$. Find the fraction (percentage) of each cycle during which $v_O > 0$, the value of V_O, the peak diode current, and the value of PIV, all for the case in which v_S is a 12-V (rms) sinusoid, $V_D \simeq 0.7$ V, and $R = 100\ \Omega$.

Ans. 97.4%; 10.1 V; 163 mA; 33.2 V

4.5.3 The Bridge Rectifier

An alternative implementation of the full-wave rectifier is shown in Fig. 4.25(a). This circuit, known as the bridge rectifier because of the similarity of its configuration to that of the Wheatstone bridge, does not require a center-tapped transformer, a distinct advantage over the full-wave rectifier circuit of Fig. 4.24. The bridge rectifier, however, requires four diodes as compared to two in the previous circuit. This is not much of a disadvantage, because diodes are inexpensive and one can buy a diode bridge in one package.

The bridge-rectifier circuit operates as follows: During the positive half-cycles of the input voltage, v_S is positive, and thus current is conducted through diode D_1, resistor R, and diode D_2. Meanwhile, diodes D_3 and D_4 will be reverse biased. Observe that there are two diodes in series in the conduction path, and thus v_O will be lower than v_S by two diode drops (compared to one drop in the circuit previously discussed). This is somewhat of a disadvantage of the bridge rectifier.

Next, consider the situation during the negative half-cycles of the input voltage. The secondary voltage v_S will be negative, and thus $-v_S$ will be positive, forcing current through D_3, R, and D_4. Meanwhile, diodes D_1 and D_2 will be reverse biased. The important point to note, though, is that during both half-cycles, current flows through R in the same direction (from right to left), and thus v_O will always be positive, as indicated in Fig. 4.25(b).

To determine the peak inverse voltage (PIV) of each diode, consider the circuit during the positive half-cycles. The reverse voltage across D_3 can be determined from the loop formed

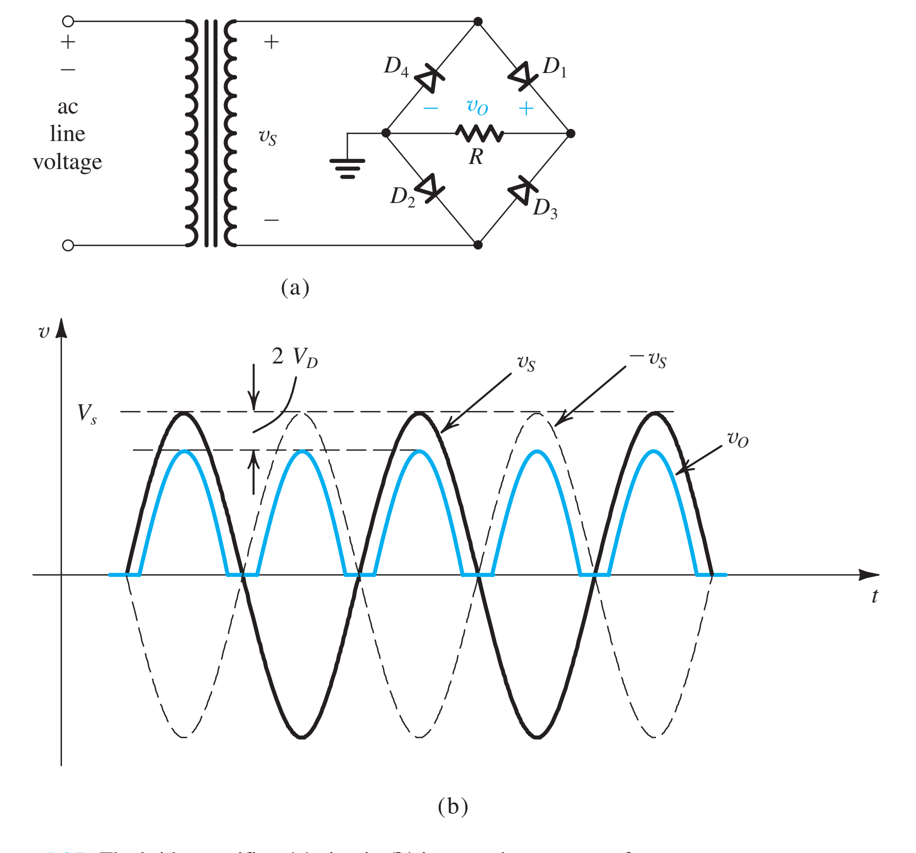


Figure 4.25 The bridge rectifier: **(a)** circuit; **(b)** input and output waveforms.

by D_3, R, and D_2 as

$$v_{D3}(\text{reverse}) = v_O + v_{D2}(\text{forward})$$

Thus the maximum value of v_{D3} occurs at the peak of v_O and is given by

$$\text{PIV} = V_s - 2V_D + V_D = V_s - V_D$$

Observe that here the PIV is about half the value for the full-wave rectifier with a center-tapped transformer. This is another advantage of the bridge rectifier.

Yet one more advantage of the bridge-rectifier circuit over that utilizing a center-tapped transformer is that only about half as many turns are required for the secondary winding of the transformer. Another way of looking at this point can be obtained by observing that each half of the secondary winding of the center-tapped transformer is utilized for only half the time. These advantages have made the bridge rectifier the most popular rectifier circuit configuration.

EXERCISE

4.21 For the bridge-rectifier circuit of Fig. 4.25(a), use the constant-voltage-drop diode model to show that (a) the average (or dc component) of the output voltage is $V_O \simeq (2/\pi)V_s - 2V_D$ and (b) the peak diode current is $(V_s - 2V_D)/R$. Find numerical values for the quantities in (a) and (b) and the PIV for the case in which v_S is a 12-V (rms) sinusoid, $V_D \simeq 0.7$ V, and $R = 100\ \Omega$.
Ans. 9.4 V; 156 mA; 16.3 V

4.5.4 The Rectifier with a Filter Capacitor—The Peak Rectifier

The pulsating nature of the output voltage produced by the rectifier circuits discussed above makes it unsuitable as a dc supply for electronic circuits. A simple way to reduce the variation of the output voltage is to place a capacitor across the load resistor. It will be shown that this **filter capacitor** serves to reduce substantially the variations in the rectifier output voltage.

To see how the rectifier circuit with a filter capacitor works, consider first the simple circuit shown in Fig. 4.26. Let the input v_I be a sinusoid with a peak value V_p, and assume the diode to be ideal. As v_I goes positive, the diode conducts and the capacitor is charged so that $v_O = v_I$. This situation continues until v_I reaches its peak value V_p. Beyond the peak, as v_I decreases, the diode becomes reverse biased and the output voltage remains constant at the value V_p. In fact, theoretically speaking, the capacitor will retain its charge and hence its voltage indefinitely, because there is no way for the capacitor to discharge. Thus the circuit provides a dc voltage output equal to the peak of the input sine wave. This is a very encouraging result in view of our desire to produce a dc output.

Next, we consider the more practical situation where a load resistance R is connected across the capacitor C, as depicted in Fig. 4.27(a). However, we will continue to assume the diode to be ideal. As before, for a sinusoidal input, the capacitor charges to the peak of the input V_p. Then the diode cuts off, and the capacitor discharges through the load resistance R. The capacitor discharge will continue for almost the entire cycle, until the time at which v_I

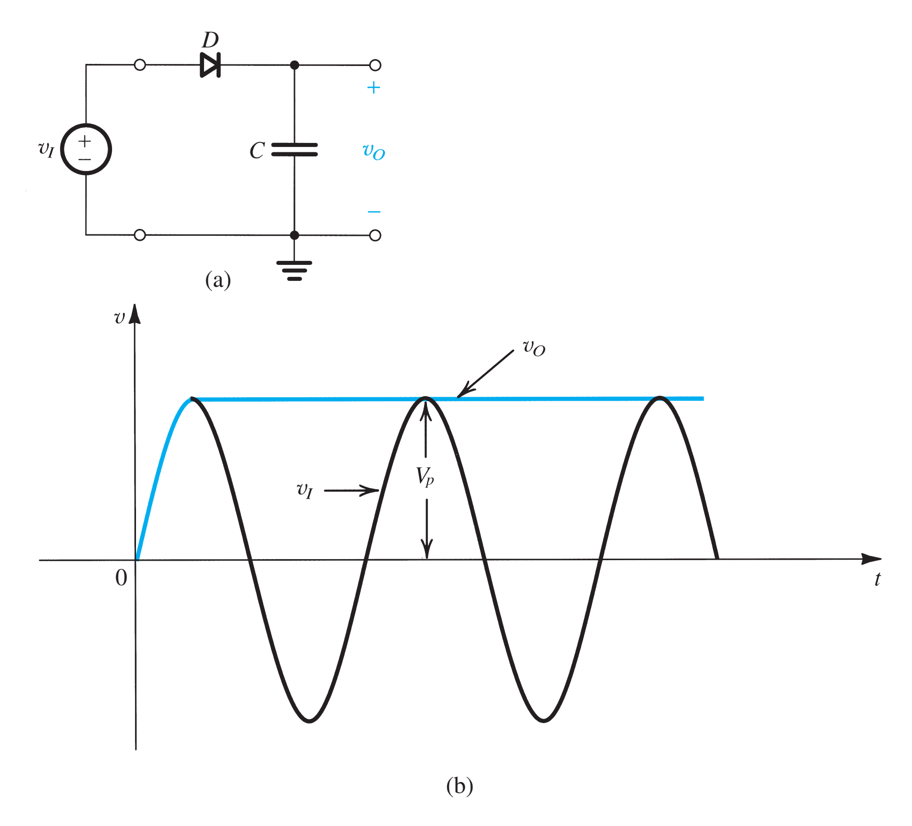


Figure 4.26 **(a)** A simple circuit used to illustrate the effect of a filter capacitor. **(b)** Input and output waveforms assuming an ideal diode. Note that the circuit provides a dc voltage equal to the peak of the input sine wave. The circuit is therefore known as a *peak rectifier* or a *peak detector*.

exceeds the capacitor voltage. Then the diode turns on again and charges the capacitor up to the peak of v_I, and the process repeats itself. Observe that to keep the output voltage from decreasing too much during capacitor discharge, one selects a value for C so that the time constant CR is much greater than the discharge interval.

We are now ready to analyze the circuit in detail. Figure 4.27(b) shows the steady-state input and output voltage waveforms under the assumption that $CR \gg T$, where T is the period of the input sinusoid. The waveforms of the load current

$$i_L = v_O/R \tag{4.23}$$

and of the diode current (when it is conducting)

$$i_D = i_C + i_L \tag{4.24}$$

$$= C\frac{dv_I}{dt} + i_L \tag{4.25}$$

are shown in Fig. 4.27(c). The following observations are in order:

1. The diode conducts for a brief interval, Δt, near the peak of the input sinusoid and supplies the capacitor with charge equal to that lost during the much longer discharge interval. The latter is approximately equal to the period T.

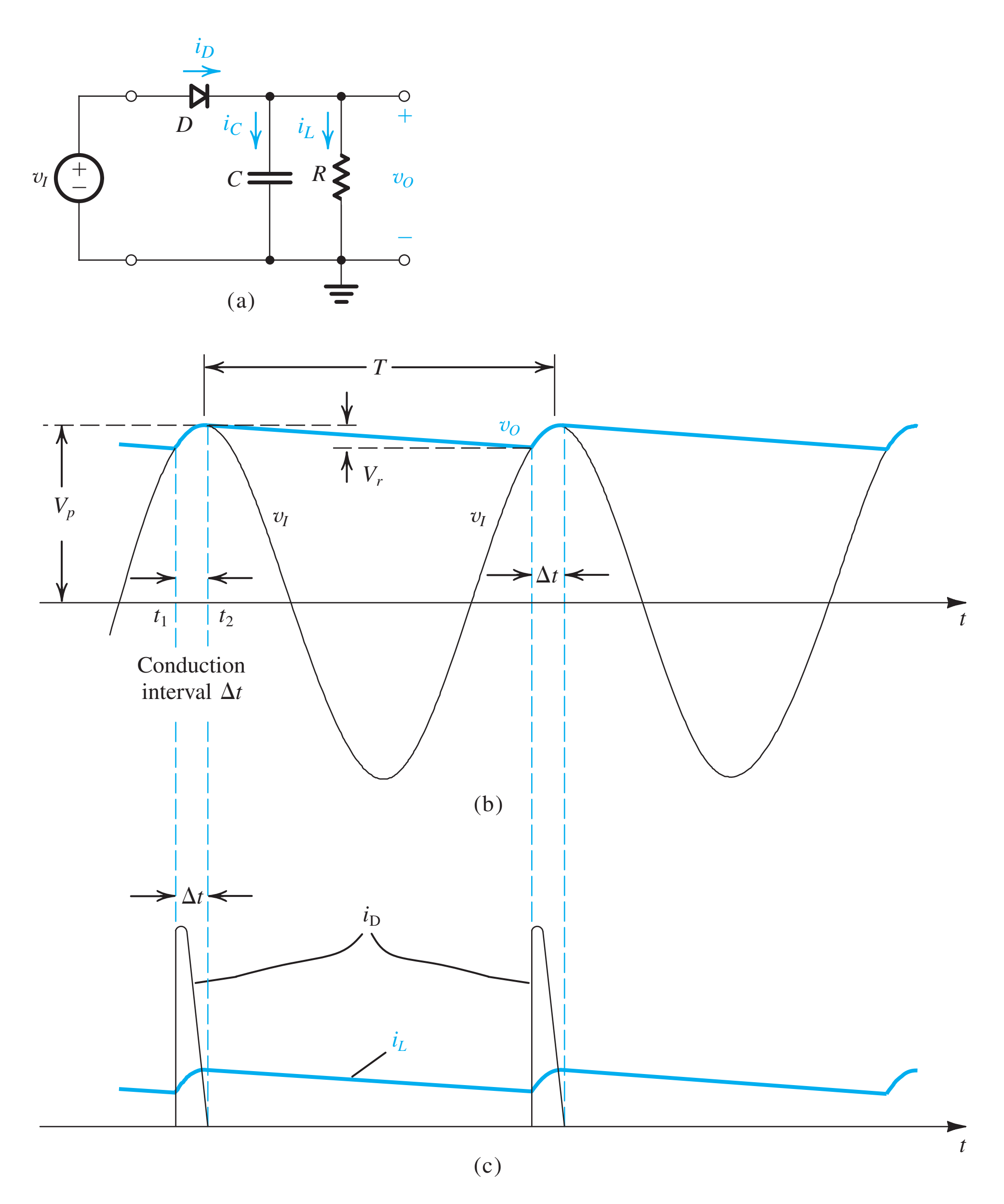


Figure 4.27 Voltage and current waveforms in the peak-rectifier circuit with $CR \gg T$. The diode is assumed ideal.

2. Assuming an ideal diode, the diode conduction begins at time t_1, at which the input v_I equals the exponentially decaying output v_O. Conduction stops at t_2 shortly after the peak of v_I; the exact value of t_2 can be determined by setting $i_D = 0$ in Eq. (4.25).
3. During the diode-off interval, the capacitor C discharges through R, and thus v_O decays exponentially with a time constant CR. The discharge interval begins just past the peak of v_I. At the end of the discharge interval, which lasts for almost the entire period T, $v_O = V_p - V_r$, where V_r is the peak-to-peak ripple voltage. When $CR \gg T$, the value of V_r is small.

4. When V_r is small, v_O is almost constant and equal to the peak value of v_I. Thus the dc output voltage is approximately equal to V_p. Similarly, the current i_L is almost constant, and its dc component I_L is given by

$$I_L = \frac{V_p}{R} \tag{4.26}$$

If desired, a more accurate expression for the output dc voltage can be obtained by taking the average of the extreme values of v_O,

$$V_O = V_p - \tfrac{1}{2}V_r \tag{4.27}$$

With these observations in hand, we now derive expressions for V_r and for the average and peak values of the diode current. During the diode-off interval, v_O can be expressed as

$$v_O = V_p e^{-t/CR}$$

At the end of the discharge interval we have

$$V_p - V_r \simeq V_p e^{-T/CR}$$

Now, since $CR \gg T$, we can use the approximation $e^{-T/CR} \simeq 1 - T/CR$ to obtain

$$V_r \simeq V_p \frac{T}{CR} \tag{4.28}$$

We observe that to keep V_r small we must select a capacitance C so that $CR \gg T$. The **ripple voltage** V_r in Eq. (4.28) can be expressed in terms of the frequency $f = 1/T$ as

$$V_r = \frac{V_p}{fCR} \tag{4.29a}$$

Using Eq. (4.26) we can express V_r by the alternate expression

$$V_r = \frac{I_L}{fC} \tag{4.29b}$$

Note that an alternative interpretation of the approximation made above is that the capacitor discharges by means of a constant current $I_L = V_p/R$. This approximation is valid as long as $V_r \ll V_p$.

Assuming that diode conduction ceases almost at the peak of v_I, we can determine the **conduction interval** Δt from

$$V_p \cos(\omega \Delta t) = V_p - V_r$$

where $\omega = 2\pi f = 2\pi/T$ is the angular frequency of v_I. Since $(\omega \Delta t)$ is a small angle, we can employ the approximation $\cos(\omega \Delta t) \simeq 1 - \frac{1}{2}(\omega \Delta t)^2$ to obtain

$$\omega \Delta t \simeq \sqrt{2V_r/V_p} \tag{4.30}$$

We note that when $V_r \ll V_p$, the conduction angle $\omega \Delta t$ will be small, as assumed.

To determine the average diode current during conduction, i_{Dav}, we equate the charge that the diode supplies to the capacitor,

$$Q_{\text{supplied}} = i_{Cav} \Delta t$$

where from Eq. (4.24),

$$i_{Cav} = i_{Dav} - I_L$$

to the charge that the capacitor loses during the discharge interval,

$$Q_{lost} = CV_r$$

to obtain, using Eqs. (4.30) and (4.29a),

$$i_{Dav} = I_L\left(1 + \pi\sqrt{2V_p/V_r}\right) \quad (4.31)$$

Observe that when $V_r \ll V_p$, the average diode current during conduction is much greater than the dc load current. This is not surprising, since the diode conducts for a very short interval and must replenish the charge lost by the capacitor during the much longer interval in which it is discharged by I_L.

The peak value of the diode current, i_{Dmax}, can be determined by evaluating the expression in Eq. (4.25) at the onset of diode conduction—that is, at $t = t_1 = -\Delta t$ (where $t = 0$ is at the peak). Assuming that i_L is almost constant at the value given by Eq. (4.26), we obtain

$$i_{Dmax} = I_L\left(1 + 2\pi\sqrt{2V_p/V_r}\right) \quad (4.32)$$

From Eqs. (4.31) and (4.32), we see that for $V_r \ll V_p$, $i_{Dmax} \simeq 2i_{Dav}$, which correlates with the fact that the waveform of i_D is almost a right-angle triangle (see Fig. 4.27c).

Example 4.8

Consider a peak rectifier fed by a 60-Hz sinusoid having a peak value $V_p = 100$ V. Let the load resistance $R = 10$ kΩ. Find the value of the capacitance C that will result in a peak-to-peak ripple of 2 V. Also, calculate the fraction of the cycle during which the diode is conducting and the average and peak values of the diode current.

Solution

From Eq. (4.29a) we obtain the value of C as

$$C = \frac{V_p}{V_r fR} = \frac{100}{2 \times 60 \times 10 \times 10^3} = 83.3\ \mu\text{F}$$

The conduction angle $\omega\Delta t$ is found from Eq. (4.30) as

$$\omega\Delta t = \sqrt{2 \times 2/100} = 0.2 \text{ rad}$$

Thus the diode conducts for $(0.2/2\pi) \times 100 = 3.18\%$ of the cycle. The average diode current is obtained from Eq. (4.31), where $I_L = 100/10 = 10$ mA, as

$$i_{Dav} = 10\left(1 + \pi\sqrt{2 \times 100/2}\right) = 324 \text{ mA}$$

The peak diode current is found using Eq. (4.32),

$$i_{Dmax} = 10\left(1 + 2\pi\sqrt{2 \times 100/2}\right) = 638 \text{ mA}$$

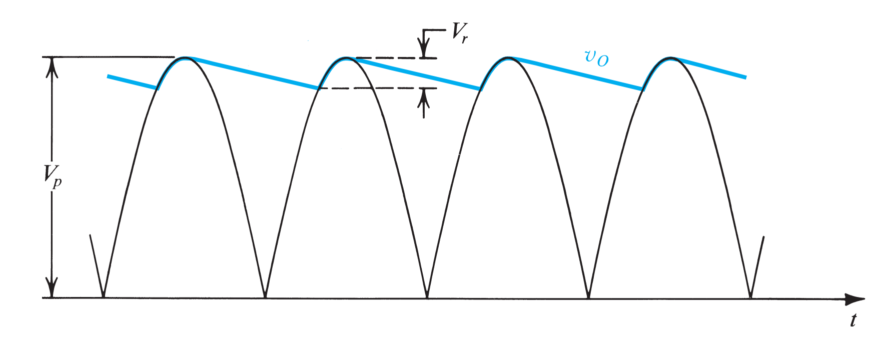


Figure 4.28 Waveforms in the full-wave peak rectifier.

The circuit of Fig. 4.27(a) is known as a half-wave **peak rectifier**. The full-wave rectifier circuits of Figs. 4.24(a) and 4.25(a) can be converted to peak rectifiers by including a capacitor across the load resistor. As in the half-wave case, the output dc voltage will be almost equal to the peak value of the input sine wave (Fig. 4.28). The ripple frequency, however, will be twice that of the input. The peak-to-peak ripple voltage, for this case, can be derived using a procedure identical to that above but with the discharge period T replaced by $T/2$, resulting in

$$V_r = \frac{V_p}{2fCR} \tag{4.33}$$

While the diode conduction interval, Δt, will still be given by Eq. (4.30), the average and peak currents in each of the diodes will be given by

$$i_{D\text{av}} = I_L\left(1 + \pi\sqrt{V_p/2V_r}\right) \tag{4.34}$$

$$i_{D\text{max}} = I_L\left(1 + 2\pi\sqrt{V_p/2V_r}\right) \tag{4.35}$$

Comparing these expressions with the corresponding ones for the half-wave case, we note that for the same values of V_p, f, R, and V_r (and thus the same I_L), we need a capacitor half the size of that required in the half-wave rectifier. Also, the current in each diode in the full-wave rectifier is approximately half that which flows in the diode of the half-wave circuit.

The analysis above assumed ideal diodes. The accuracy of the results can be improved by taking the diode voltage drop into account. This can be easily done by replacing the peak voltage V_p to which the capacitor charges with $(V_p - V_D)$ for the half-wave circuit and the full-wave circuit using a center-tapped transformer and with $(V_p - 2V_D)$ for the bridge-rectifier case.

We conclude this section by noting that peak-rectifier circuits find application in signal-processing systems where it is required to detect the peak of an input signal. In such a case, the circuit is referred to as a **peak detector**. A particularly popular application of the peak detector is in the design of a demodulator for amplitude-modulated (AM) signals. We shall not discuss this application further here.

EXERCISES

4.22 Derive the expressions in Eqs. (4.33), (4.34), and (4.35).

4.23 Consider a bridge-rectifier circuit with a filter capacitor C placed across the load resistor R for the case in which the transformer secondary delivers a sinusoid of 12 V (rms) having a 60-Hz frequency and assuming $V_D = 0.8$ V and a load resistance $R = 100\ \Omega$. Find the value of C that results in a ripple voltage no larger than 1 V peak-to-peak. What is the dc voltage at the output? Find the load current. Find the diodes' conduction angle. Provide the average and peak diode currents. What is the peak reverse voltage across each diode? Specify the diode in terms of its peak current and its PIV.
Ans. 1281 μF; 15.4 V or (a better estimate) 14.9 V; 0.15 A; 0.36 rad (20.7°); 1.45 A; 2.74 A; 16.2 V. Thus select a diode with 3.5-A to 4-A peak current and a 20-V PIV rating.

THE EARLIEST SEMICONDUCTOR DIODE:

The cat's whisker or crystal detector was the first electronic diode to be commercialized as an envelope detector for the radio-frequency signals used in radio telephony. The earliest diode, invented in Germany by Karl Ferdinand Braun, consisted of a small slab of galena (lead sulfide) to which contact was made by sharpened spring wire, which could be adjusted. For this and other contributions to early radios, Braun received the Nobel Prize in Physics in 1909. The silicon-based point-contact diode, later refined and packaged, was an important solid-state component of radar equipment during World War II.

4.5.5 Precision Half-Wave Rectifier—The Superdiode[4]

The rectifier circuits studied thus far suffer from having one or two diode drops in the signal paths. Thus these circuits work well only when the signal to be rectified is much larger than the voltage drop of a conducting diode (0.7 V or so). In such a case, the details of the diode forward characteristics or the exact value of the diode voltage do not play a prominent role in determining circuit performance. This is indeed the case in the application of rectifier circuits in power-supply design. There are other applications, however, where the signal to be rectified is small (e.g., on the order of 100 mV or so) and thus clearly insufficient to turn on a diode. Also, in instrumentation applications, the need arises for rectifier circuits with very precise and predictable transfer characteristics. For these applications, a class of circuits has been developed utilizing op amps (Chapter 2) together with diodes to provide precision rectification. In the following discussion, we study one such circuit. A comprehensive study of op amp–diode circuits is available on the website.

[4]This section requires knowledge of operational amplifiers (Chapter 2).

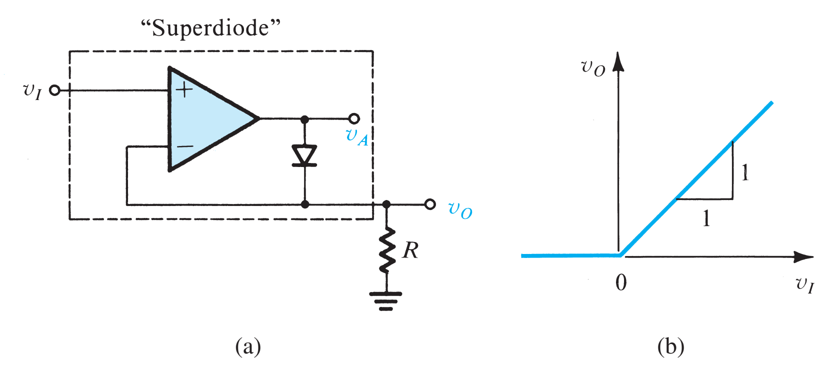


Figure 4.29 **(a)** The "superdiode" precision half-wave rectifier and **(b)** its almost-ideal transfer characteristic. Note that when $v_I > 0$ and the diode conducts, the op amp supplies the load current, and the source is conveniently buffered, an added advantage. Not shown are the op-amp power supplies.

Figure 4.29(a) shows a precision half-wave rectifier circuit consisting of a diode placed in the negative-feedback path of an op amp, with R being the rectifier load resistance. The op amp, of course, needs power supplies for its operation. For simplicity, these are not shown in the circuit diagram. The circuit works as follows: If v_I goes positive, the output voltage v_A of the op amp will go positive and the diode will conduct, thus establishing a closed feedback path between the op amp's output terminal and the negative input terminal. This negative-feedback path will cause a virtual short circuit to appear between the two input terminals of the op amp. Thus the voltage at the negative input terminal, which is also the output voltage v_O, will equal (to within a few millivolts) that at the positive input terminal, which is the input voltage v_I,

$$v_O = v_I \quad v_I \geq 0$$

Note that the offset voltage ($\simeq 0.7$ V) exhibited in the simple half-wave rectifier circuit of Fig. 4.23 is no longer present. For the op-amp circuit to start operation, v_I has to exceed only a negligibly small voltage equal to the diode drop divided by the op amp's open-loop gain. In other words, the straight-line transfer characteristic v_O–v_I almost passes through the origin. This makes this circuit suitable for applications involving very small signals.

Consider now the case when v_I goes negative. The op amp's output voltage v_A will tend to follow and go negative. This will reverse-bias the diode, and no current will flow through resistance R, causing v_O to remain equal to 0 V. Thus, for $v_I < 0$, $v_O = 0$. Since in this case the diode is off, the op amp will be operating in an open-loop fashion, and its output will be at its negative saturation level.

The transfer characteristic of this circuit will be that shown in Fig. 4.29(b), which is almost identical to the ideal characteristic of a half-wave rectifier. The nonideal diode characteristics have been almost completely masked by placing the diode in the negative-feedback path of an op amp. This is another dramatic application of negative feedback, a subject we will study formally in Chapter 11. The combination of diode and op amp, shown in the dashed box in Fig. 4.29(a), is appropriately referred to as a "superdiode."

EXERCISES

4.24 Consider the **operational rectifier** or superdiode circuit of Fig. 4.29(a), with $R = 1$ kΩ. For $v_I =$ 10 mV, 1 V, and –1 V, what are the voltages that result at the rectifier output and at the output of the op amp? Assume that the op amp is ideal and that its output saturates at ±12 V. The diode has a 0.7-V drop at 1-mA current.
Ans. 10 mV, 0.59 V; 1 V, 1.7 V; 0 V, –12 V

4.25 If the diode in the circuit of Fig. 4.29(a) is reversed, find the transfer characteristic v_O as a function of v_I.
Ans. $v_O = 0$ for $v_I \geq 0$; $v_O = v_I$ for $v_I \leq 0$

4.6 Limiting and Clamping Circuits



In this section, we shall present additional nonlinear circuit applications of diodes.

4.6.1 Limiter Circuits

Figure 4.30 shows the general transfer characteristic of a limiter circuit. As indicated, for inputs in a certain range, $L_-/K \leq v_I \leq L_+/K$, the limiter acts as a linear circuit, providing an output proportional to the input, $v_O = Kv_I$. Although in general K can be greater than 1, the

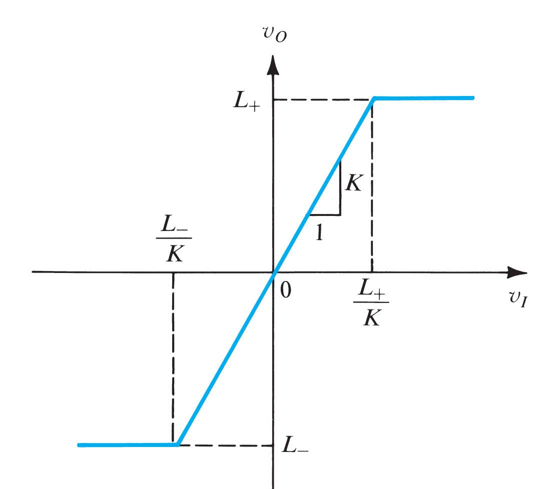


Figure 4.30 General transfer characteristic for a limiter circuit.

circuits discussed in this section have $K \leq 1$ and are known as **passive limiters**. (Examples of active limiters will be presented in Chapter 18.) If v_I exceeds the upper *threshold* (L_+/K), the output voltage is *limited* or clamped to the upper limiting level L_+. On the other hand, if v_I is reduced below the lower limiting threshold (L_-/K), the output voltage v_O is limited to the lower limiting level L_-.

The general transfer characteristic of Fig. 4.30 describes a **double limiter**—that is, a limiter that works on both the positive and negative peaks of an input waveform. **Single limiters**, of course, exist. Finally, note that if an input waveform such as that shown in Fig. 4.31 is fed to a double limiter, its two peaks will be *clipped off*. Limiters therefore are sometimes referred to as **clippers**.

The limiter whose characteristics are depicted in Fig. 4.30 is described as a **hard limiter**. **Soft limiting** is characterized by smoother transitions between the linear region and the saturation regions and a slope greater than zero in the saturation regions, as illustrated in Fig. 4.32. Depending on the application, either hard or soft limiting may be preferred.

Limiters find application in a variety of signal-processing systems. One of their simplest applications is in limiting the voltage between the two input terminals of an op amp to a value lower than the breakdown voltage of the transistors that make up the input stage of the op-amp circuit. We will have more to say on this and other limiter applications at later points in this book.

Diodes can be combined with resistors to provide simple realizations of the limiter function. A number of examples are depicted in Fig. 4.33. In each part of the figure both the circuit and its transfer characteristic are given. The transfer characteristics are obtained using the constant-voltage-drop ($V_D = 0.7$ V) diode model but assuming a smooth transition between the linear and saturation regions of the transfer characteristic.

The circuit in Fig. 4.33(a) is that of the half-wave rectifier except that here the output is taken across the diode. For $v_I < 0.5$ V, the diode is cut off, no current flows, and the voltage drop across R is zero; thus $v_O = v_I$. As v_I exceeds 0.5 V, the diode turns on, eventually limiting

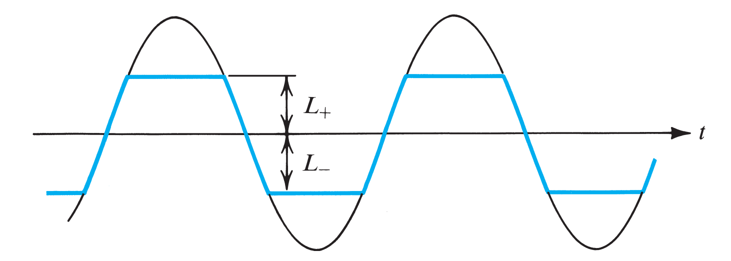


Figure 4.31 Applying a sine wave to a limiter can result in clipping off its two peaks.

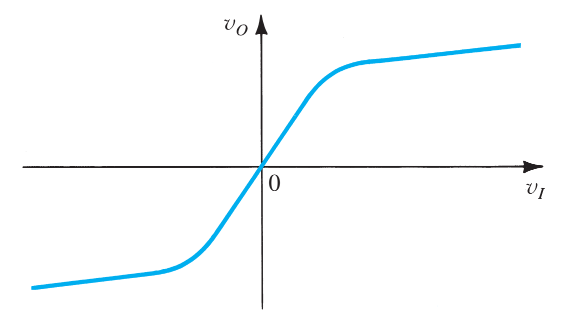


Figure 4.32 Soft limiting.

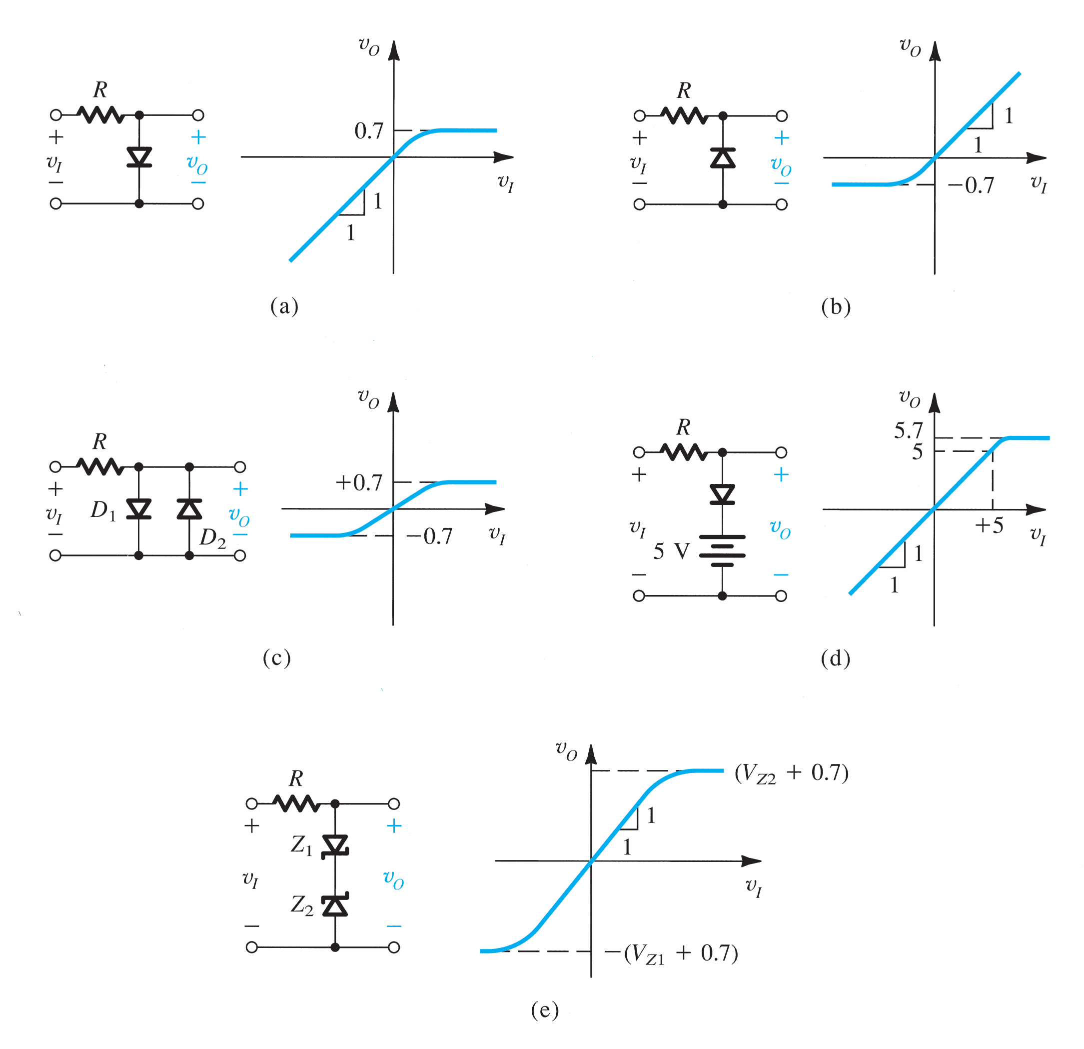


Figure 4.33 A variety of basic limiting circuits.

v_O to one diode drop (0.7 V). The circuit of Fig. 4.33(b) is similar to that in Fig. 4.33(a) except that the diode is reversed.

Double limiting can be implemented by placing two diodes of opposite polarity in parallel, as shown in Fig. 4.33(c). Here the linear region of the characteristic is obtained for $-0.5\text{V} \le v_I \le 0.5$ V. For this range of v_I, both diodes are off and $v_O = v_I$. As v_I exceeds 0.5 V, D_1 turns on and eventually limits v_O to +0.7 V. Similarly, as v_I goes more negative than –0.5 V, D_2 turns on and eventually limits v_O to –0.7 V.

The thresholds and saturation levels of diode limiters can be controlled by using strings of diodes and/or by connecting a dc voltage in series with the diode(s). The latter idea is illustrated in Fig. 4.33(d). Finally, rather than strings of diodes, we may use two zener diodes in series, as shown in Fig. 4.33(e). In this circuit, limiting occurs in the positive direction at a voltage of $V_{Z2} + 0.7$, where 0.7 V represents the voltage drop across zener diode Z_1 when conducting in the *forward* direction. For negative inputs, Z_1 acts as a zener, while Z_2 conducts

in the forward direction. It should be mentioned that pairs of zener diodes connected in series are available commercially for applications of this type under the name **double-anode zener**.

More flexible limiter circuits are possible if op amps are combined with diodes and resistors. Examples of such circuits are discussed in Chapter 18.

EXERCISE

4.26 Assuming the diodes to be ideal, describe the transfer characteristic of the circuit shown in Fig. E4.26.

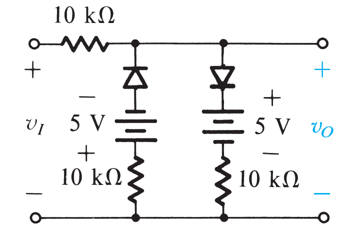


Figure E4.26

Ans.

$v_O = v_I$	for $-5 \le v_I \le +5$
$v_O = \frac{1}{2}v_I - 2.5$	for $v_I \le -5$
$v_O = \frac{1}{2}v_I + 2.5$	for $v_I \ge +5$

4.6.2 The Clamped Capacitor or DC Restorer

If in the basic peak-rectifier circuit, the output is taken across the diode rather than across the capacitor, an interesting circuit with important applications results. The circuit, called a dc restorer, is shown in Fig. 4.34 fed with a square wave. Because of the polarity in which the diode is connected, the capacitor will charge to a voltage v_C with the polarity indicated in Fig. 4.34 and equal to the magnitude of the most negative peak of the input signal. Subsequently, the diode turns off and the capacitor retains its voltage indefinitely. If,

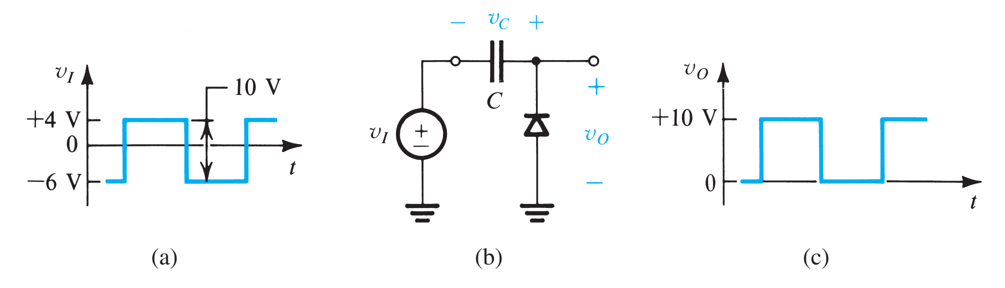


Figure 4.34 The clamped capacitor or dc restorer with a square-wave input and no load.

for instance, the input square wave has the arbitrary levels –6 V and +4 V, then v_C will be equal to 6 V. Now, since the output voltage v_O is given by

$$v_O = v_I + v_C$$

it follows that the output waveform will be identical to that of the input, except that it is shifted upward by v_C volts. In our example the output will thus be a square wave with levels of 0 V and +10 V.

Another way of visualizing the operation of the circuit in Fig. 4.34 is to note that because the diode is connected across the output with the polarity shown, it prevents the output voltage from going below 0 V (by conducting and charging up the capacitor, thus causing the output to rise to 0 V), but this connection will not constrain the positive excursion of v_O. The output waveform will therefore have its lowest peak *clamped* to 0 V, which is why the circuit is called a **clamped capacitor**. It should be obvious that reversing the diode polarity will provide an output waveform whose highest peak is clamped to 0 V. In either case, the output waveform will have a finite average value or dc component. This dc component is entirely unrelated to the average value of the input waveform. As an application, consider a pulse signal being transmitted through a capacitively coupled or ac-coupled system. The capacitive coupling will cause the pulse train to lose whatever dc component it originally had. Feeding the resulting pulse waveform to a clamping circuit provides it with a well-determined dc component, a process known as **dc restoration**. This is why the circuit is also called a **dc restorer**.

Restoring dc is useful because the dc component or average value of a pulse waveform is an effective measure of its **duty cycle**.[5] The duty cycle of a pulse waveform can be modulated (in a process called **pulsewidth modulation**) and made to carry information. In such a system, detection or demodulation could be achieved simply by feeding the received pulse waveform to a dc restorer and then using a simple *RC* low-pass filter to separate the average of the output waveform from the superimposed pulses.

When a load resistance *R* is connected across the diode in a clamping circuit, as shown in Fig. 4.35, the situation changes significantly. While the output is above ground, a current must flow in *R*. Since at this time the diode is off, this current obviously comes from the capacitor,

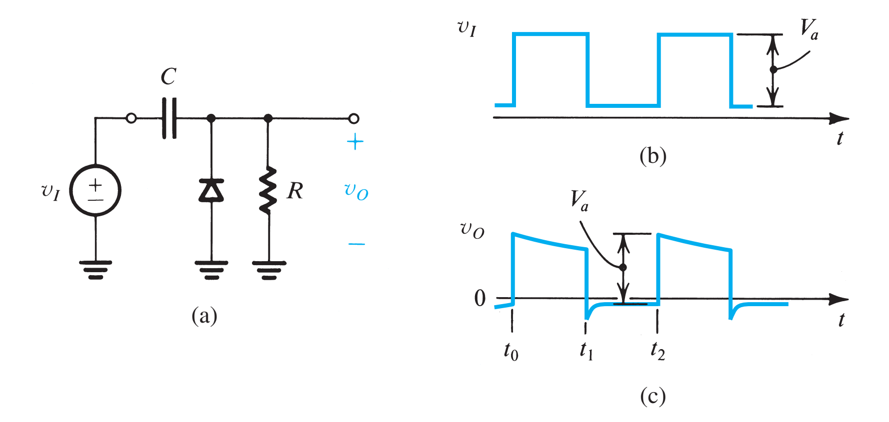


Figure 4.35 The clamped capacitor with a load resistance *R*.

[5]The duty cycle of a pulse waveform is the proportion of each cycle occupied by the pulse. In other words, it is the pulse width expressed as a fraction of the pulse period.

thus causing the capacitor to discharge and the output voltage to fall. This is shown in Fig. 4.35 for a square-wave input. During the interval t_0 to t_1, the output voltage falls exponentially with time constant CR. At t_1 the input decreases by V_a volts, and the output attempts to follow. This causes the diode to conduct heavily and to quickly charge the capacitor. At the end of the interval t_1 to t_2, the output voltage would normally be a few tenths of a volt negative (e.g., –0.5 V). Then, as the input rises by V_a volts (at t_2), the output follows, and the cycle repeats itself. In the steady state the charge lost by the capacitor during the interval t_0 to t_1 is recovered during the interval t_1 to t_2. This charge equilibrium enables us to calculate the average diode current as well as the details of the output waveform.

4.6.3 The Voltage Doubler

Figure 4.36(a) shows a circuit composed of two sections in cascade: a clamped capacitor formed by C_1 and D_1, and a peak rectifier formed by D_2 and C_2. When excited by a sinusoid of amplitude V_p the clamping section provides the voltage waveform v_{D1} shown, assuming ideal diodes, in Fig. 4.36(b). Note that while the positive peaks are clamped to 0 V, the negative peak reaches $-2V_p$. In response to this waveform, the peak-detector section provides across capacitor C_2 a dc voltage equal to the negative peak of v_{D1}, that is, $-2V_p$. Because the output voltage is double the input peak, the circuit is known as a voltage doubler. The technique can be extended to provide output dc voltages that are higher multiples of V_p.

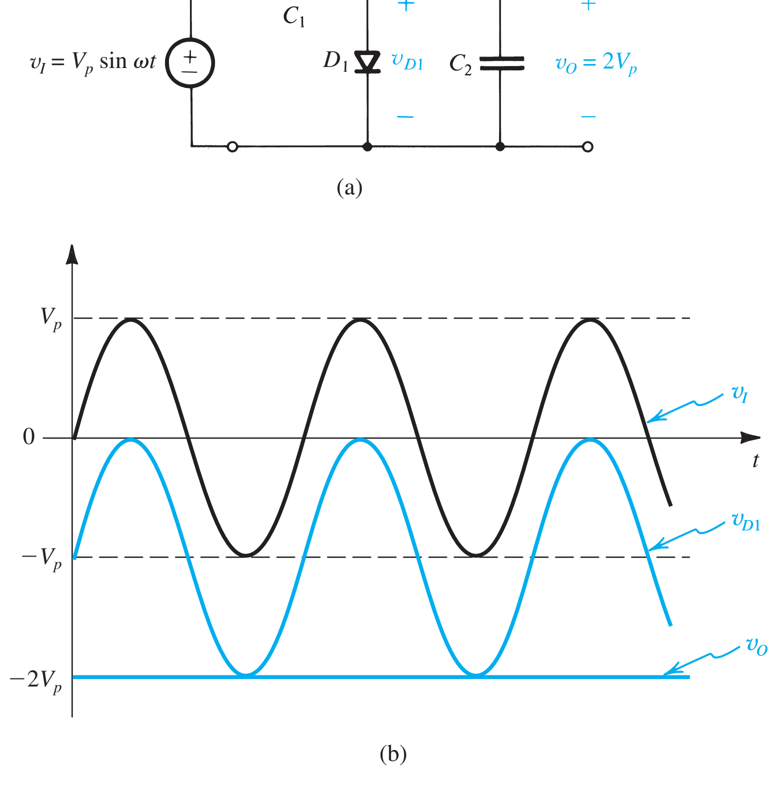


Figure 4.36 Voltage doubler: **(a)** circuit; **(b)** waveforms of the input voltage, the voltage across D_1, and the output voltage $v_o = -2V_p$.

EXERCISE

4.27 If the diode in the circuit of Fig. 4.34 is reversed, what will the dc component of v_O become?
Ans. –5 V

4.7 Special Diode Types



In this section, we discuss briefly some important special types of diodes.

4.7.1 The Schottky-Barrier Diode (SBD)

The Schottky-barrier diode (SBD) is formed by bringing metal into contact with a moderately doped *n*-type semiconductor material. The resulting metal–semiconductor junction behaves like a diode, conducting current in one direction (from the metal anode to the semiconductor cathode) and acting as an open circuit in the other, and is known as the Schottky-barrier diode or simply the Schottky diode. In fact, the current–voltage characteristic of the SBD is remarkably similar to that of a *pn*-junction diode, with two important exceptions:

1. In the SBD, current is conducted by majority carriers (electrons). Thus the SBD does not exhibit the minority-carrier charge-storage effects found in forward-biased *pn* junctions. As a result, Schottky diodes can be switched from on to off, and vice versa, much faster than is possible with *pn*-junction diodes.
2. The forward voltage drop of a conducting SBD is lower than that of a *pn*-junction diode. For example, an SBD made of silicon exhibits a forward voltage drop of 0.3 V to 0.5 V, compared to the 0.6 V to 0.8 V found in silicon *pn*-junction diodes. SBDs can also be made of gallium arsenide (GaAs) and, in fact, play an important role in the design of GaAs circuits.[6] Gallium-arsenide SBDs exhibit forward voltage drops of about 0.7 V.

Apart from GaAs circuits, Schottky diodes find application in the design of a special form of bipolar-transistor logic circuits, known as Schottky-TTL, where TTL stands for transistor–transistor logic.

Before leaving the subject of Schottky-barrier diodes, it is important to note that not every metal–semiconductor contact is a diode. In fact, metal is commonly deposited on the semiconductor surface in order to make terminals for the semiconductor devices and to connect different devices in an integrated-circuit chip. Such metal–semiconductor contacts are known as **ohmic contacts** to distinguish them from the rectifying contacts that result in SBDs. Ohmic contacts are usually made by depositing metal on very heavily doped (and thus low-resistivity) semiconductor regions. (Recall that SBDs use moderately doped material.)

[6]The website accompanying this text contains material on GaAs circuits.

4.7.2 Varactors

In Chapter 3 we learned that reverse-biased *pn* junctions exhibit a charge-storage effect that is modeled with the depletion-layer or junction capacitance C_j. As Eq. (3.49) indicates, C_j is a function of the reverse-bias voltage V_R. This dependence turns out to be useful in a number of applications, such as the automatic tuning of radio receivers. Special diodes are therefore fabricated to be used as voltage-variable capacitors known as **varactors**. These devices are optimized to make the capacitance a strong function of voltage by arranging that the grading coefficient m is 3 or 4.

4.7.3 Photodiodes

If a reverse-biased *pn* junction is illuminated—that is, exposed to incident light—the photons impacting the junction cause covalent bonds to break, and thus electron-hole pairs are generated in the depletion layer. The electric field in the depletion region then sweeps the liberated electrons to the n side and the holes to the p side, giving rise to a reverse current across the junction. This current, known as photocurrent, is proportional to the intensity of the incident light. Such a diode, called a photodiode, can be used to convert light signals into electrical signals.

Photodiodes are usually fabricated using a compound semiconductor[7] such as gallium arsenide. The photodiode is an important component of a growing family of circuits known as **optoelectronics** or **photonics**. As the name implies, such circuits utilize an optimum combination of electronics and optics for signal processing, storage, and transmission. Usually, electronics is the preferred means for signal processing, whereas optics is most suited for transmission and storage. Examples include fiber-optic transmission of telephone and television signals and the use of optical storage in CD-ROM computer discs. Optical transmission provides very wide bandwidths and low signal attenuation. Optical storage allows vast amounts of data to be stored reliably in a small space.

Finally, we should note that without reverse bias, the illuminated photodiode functions as a **solar cell**. Usually fabricated from low-cost silicon, a solar cell converts light to electrical energy.

4.7.4 Light-Emitting Diodes (LEDs)

The light-emitting diode (LED) performs the inverse of the function of the photodiode; it converts a forward current into light. The reader will recall from Chapter 3 that in a forward-biased *pn* junction, minority carriers are injected across the junction and diffuse into the p and n regions. The diffusing minority carriers then recombine with the majority carriers. Such recombination can be made to give rise to light emission. This can be done by fabricating the *pn* junction using a semiconductor of the type known as direct-bandgap materials. Gallium arsenide belongs to this group and can thus be used to fabricate light-emitting diodes.

The light emitted by an LED is proportional to the number of recombinations that take place, which in turn is proportional to the forward current in the diode.

[7]Whereas an elemental semiconductor, such as silicon, uses an element from column IV of the periodic table, a compound semiconductor uses a combination of elements from columns III and V or II and VI. For example, GaAs is formed of gallium (column III) and arsenic (column V) and is thus known as a III-V compound.

LEDs are very popular devices. They find application in the design of numerous types of displays, including the displays of laboratory instruments such as digital voltmeters. They can be made to produce light in a variety of colors. Furthermore, LEDs can be designed so as to produce coherent light with a very narrow bandwidth. The resulting device is a **laser diode**. Laser diodes find application in optical communication systems and in DVD players, among other things.

Combining an LED with a photodiode in the same package results in a device known as an **optoisolator**. The LED converts an electrical signal applied to the optoisolator into light, which the photodiode detects and converts back to an electrical signal at the output of the optoisolator. Use of the optoisolator provides complete electrical isolation between the electrical circuit that is connected to the isolator's input and the circuit that is connected to its output. Such isolation can be useful in reducing the effect of electrical interference on signal transmission within a system, and thus optoisolators are frequently employed in the design of digital systems. They can also be used in the design of medical instruments to reduce the risk of electrical shock to patients.

Note that the optical coupling between an LED and a photodiode need not be accomplished inside a small package. Indeed, it can be implemented over a long distance using an optical fiber, as is done in fiber-optic communication links.

FROM INDICATION TO ILLUMINATION:

Light-emitting diodes (LEDs), which once served only as low-powered status indicators, are now lighting our way! Increasingly, automotive lighting uses LEDs; increasingly, too, LED bulbs of higher and higher power are replacing both incandescent and fluorescent lighting in homes and offices. Incandescent bulbs are only 5% efficient in the conversion of electricity into light—the other 95% is dissipated as heat. The light conversion efficiency of LEDs, however, is 60%. Moreover, LEDs last 25 times longer (25,000 hours) than incandescent bulbs and 3 times longer than fluorescents.

Summary

- In the forward direction, the ideal diode conducts any current forced by the external circuit while displaying a zero voltage drop. The ideal diode does not conduct in the reverse direction; any applied voltage appears as reverse bias across the diode.
- The unidirectional-current-flow property makes the diode useful in the design of rectifier circuits.
- The forward conduction of practical silicon-junction diodes is accurately characterized by the relationship $i = I_S e^{v/V_T}$.
- A silicon diode conducts a negligible current until the forward voltage is at least 0.5 V. Then the current increases rapidly, with the voltage drop increasing by 60 mV for every decade of current change.
- In the reverse direction, a silicon diode conducts a current on the order of 10^{-9} A. This current is much greater than I_S because of leakage effects and increases with the magnitude of reverse voltage.
- Beyond a certain value of reverse voltage (that depends on the diode), breakdown occurs, and current increases rapidly with a small corresponding increase in voltage.
- Diodes designed to operate in the breakdown region are called zener diodes. They are employed in the design of voltage regulators whose function is to provide a constant dc voltage that varies little with variations in power-supply voltage and/or load current.

- In many applications, a conducting diode is modeled as having a constant voltage drop, usually approximately 0.7 V.
- A diode biased to operate at a dc current I_D has a small-signal resistance $r_d = V_T/I_D$.
- Rectifiers convert ac voltages into unipolar voltages. Half-wave rectifiers do this by passing the voltage in half of each cycle and blocking the opposite-polarity voltage in the other half of the cycle. Full-wave rectifiers accomplish the task by passing the voltage in half of each cycle and inverting the voltage in the other half-cycle.
- The bridge-rectifier circuit is the preferred full-wave rectifier configuration.
- The variation of the output waveform of the rectifier is reduced considerably by connecting a capacitor C across the output load resistance R. The resulting circuit is the peak rectifier. The output waveform then consists of a dc voltage almost equal to the peak of the input sine wave, V_p, on which is superimposed a ripple component of frequency $2f$ (in the full-wave case) and of peak-to-peak amplitude $V_r = V_p/2fCR$. To reduce this ripple voltage further, a voltage regulator is employed.
- Combination of diodes, resistors, and possibly reference voltages can be used to design voltage limiters that prevent one or both extremities of the output waveform from going beyond predetermined values, the limiting level(s).
- Applying a time-varying waveform to a circuit consisting of a capacitor in series with a diode and taking the output across the diode provides a clamping function. Specifically, depending on the polarity of the diode, either the positive or negative peaks of the signal will be clamped to the voltage at the other terminal of the diode (usually ground). In this way the output waveform has a nonzero average or dc component, and the circuit is known as a dc restorer.
- By cascading a clamping circuit with a peak-rectifier circuit, a voltage doubler is realized.

PROBLEMS

Computer Simulation Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as allowable signal swing and amplifier nonlinear distortion. Instructions to assist in setting up PSPice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 4.1: The Ideal Diode

4.1 An AA flashlight cell, whose Thévenin equivalent is a voltage source of 1.5 V and a resistance of 1 Ω, is connected to the terminals of an ideal diode. Describe two possible situations that result. What are the diode current and terminal voltage when (a) the connection is between the diode cathode and the positive terminal of the battery and (b) the anode and the positive terminal are connected?

4.2 For the circuits shown in Fig. P4.2 using ideal diodes, find the values of the voltages and currents indicated.

4.3 For the circuits shown in Fig. P4.3 using ideal diodes, find the values of the labeled voltages and currents.

4.4 In each of the ideal-diode circuits shown in Fig. P4.4, v_I is a 1-kHz, 5-V peak sine wave. Sketch the waveform resulting at v_O. What are its positive and negative peak values?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

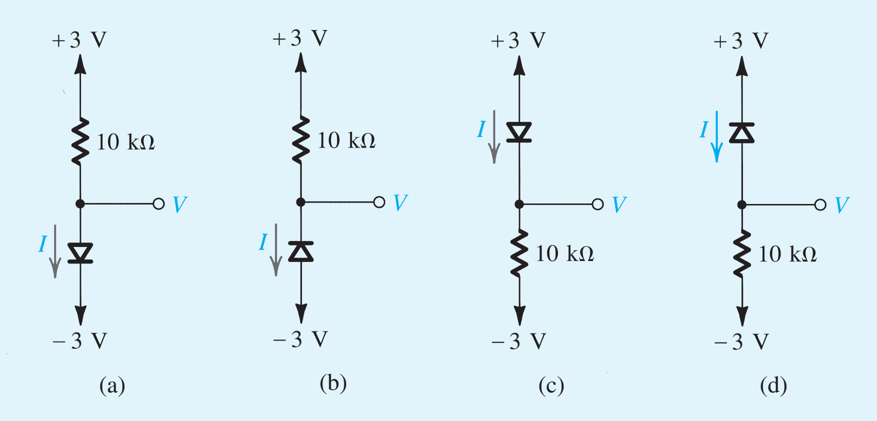


Figure P4.2

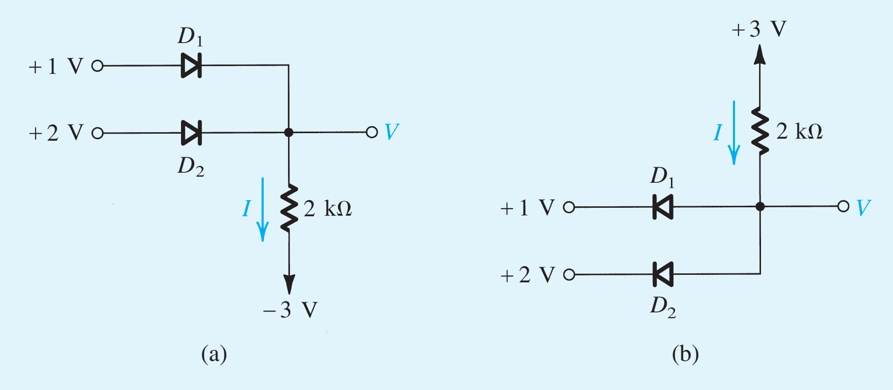


Figure P4.3

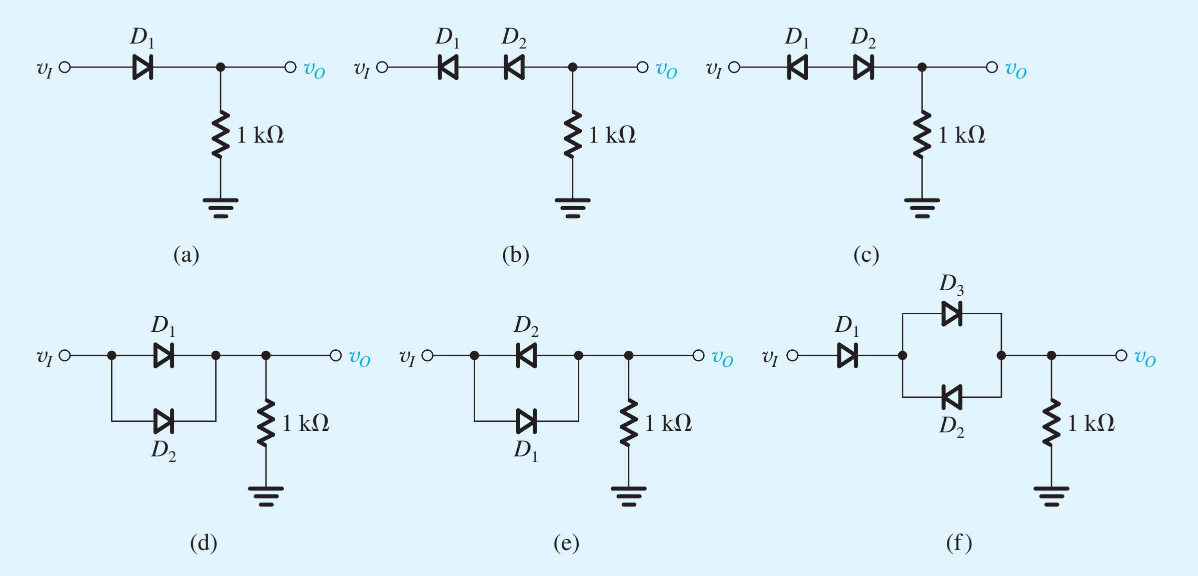


Figure P4.4

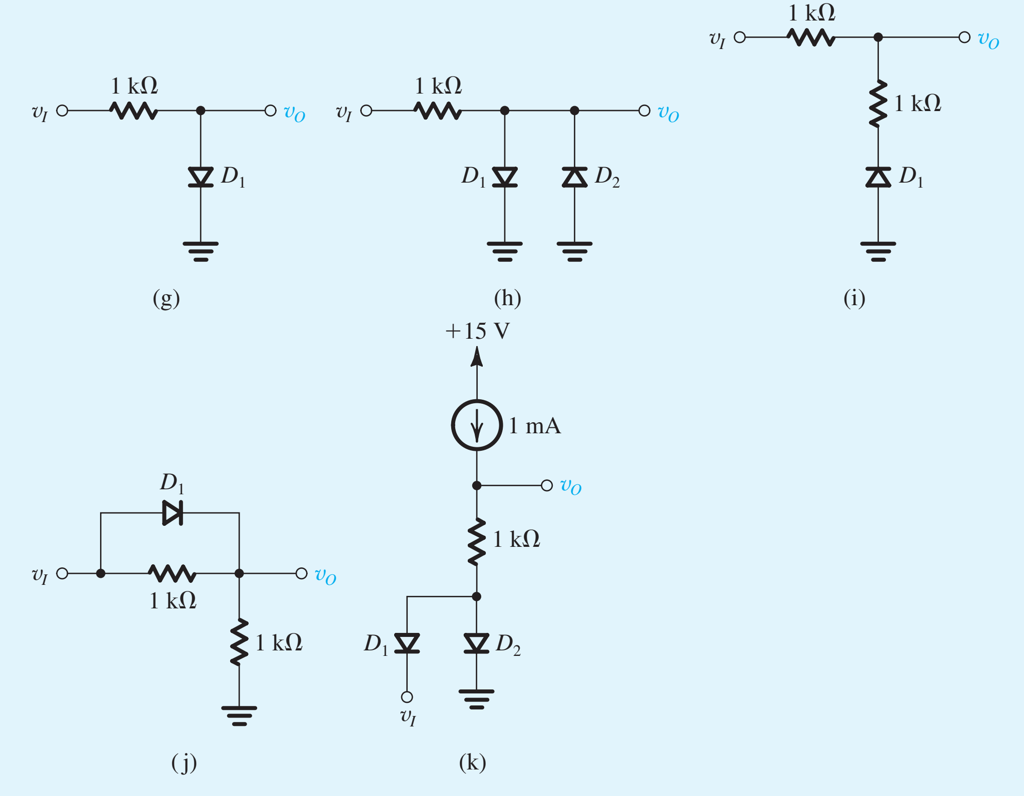


Figure P4.4 *continued*

4.5 The circuit shown in Fig. P4.5 is a model for a battery charger. Here v_I is a 6-V peak sine wave, D_1 and D_2 are ideal diodes, I is a 60-mA current source, and B is a 3-V battery. Sketch and label the waveform of the battery current i_B. What is its peak value? What is its average value? If the peak value of v_I is reduced by 10%, what do the peak and average values of i_B become?

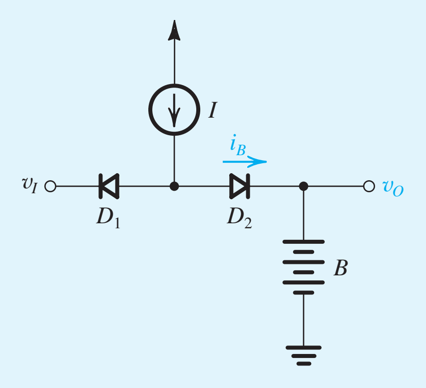


Figure P4.5

4.6 The circuits shown in Fig. P4.6 can function as logic gates for input voltages that are either high or low. Using "1" to denote the high value and "0" to denote the low value, prepare a table with four columns including all possible input combinations and the resulting values of X and Y. What logic function is X of A and B? What logic function is Y of A and B? For what values of A and B do X and Y have the same value? For what values of A and B do X and Y have opposite values?

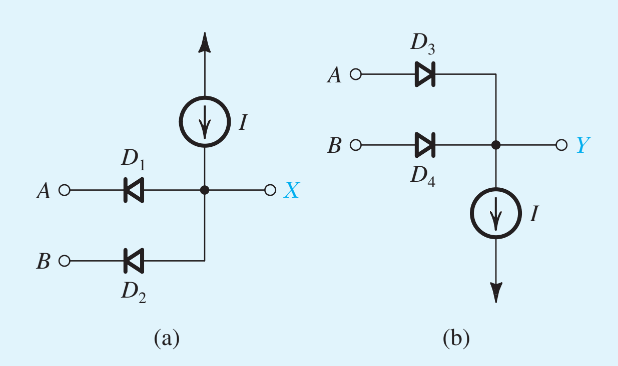


Figure P4.6

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

D 4.7 For the logic gate of Fig. 4.5(a), assume ideal diodes and input voltage levels of 0 V and +5 V. Find a suitable value for R so that the current required from each of the input signal sources does not exceed 0.2 mA.

D 4.8 Repeat Problem 4.7 for the logic gate of Fig. 4.5(b).

4.9 Assuming that the diodes in the circuits of Fig. P4.9 are ideal, find the values of the labeled voltages and currents.

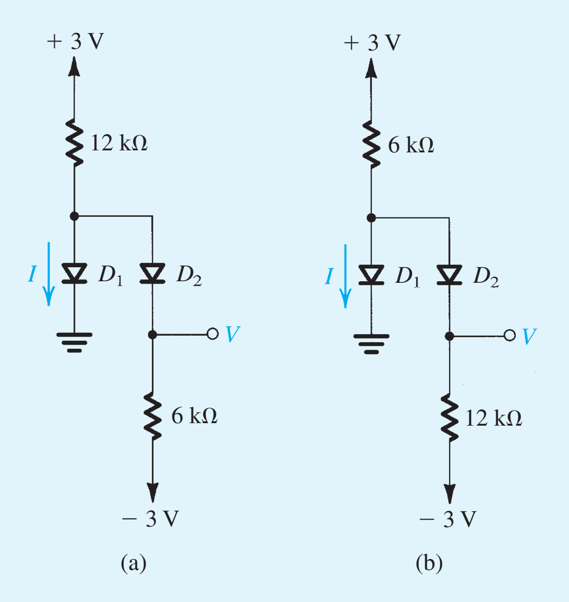


Figure P4.9

4.10 Assuming that the diodes in the circuits of Fig. P4.10 are ideal, utilize Thévenin's theorem to simplify the circuits and thus find the values of the labeled currents and voltages.

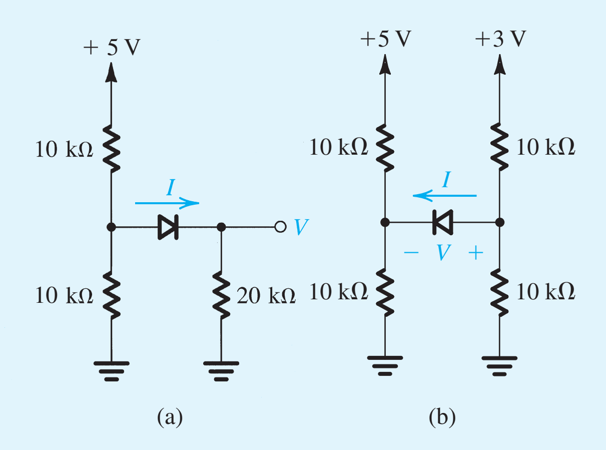


Figure P4.10

D 4.11 For the rectifier circuit of Fig. 4.3(a), let the input sine wave have 120-V rms value and assume the diode to be ideal. Select a suitable value for R so that the peak diode current does not exceed 40 mA. What is the greatest reverse voltage that will appear across the diode?

4.12 Consider the rectifier circuit of Fig. 4.3(a) in the event that the input source v_I has a source resistance R_s. For the case $R_s = R$ and assuming the diode to be ideal, sketch and clearly label the transfer characteristic v_O versus v_I.

4.13 A symmetrical square wave of 5-V peak-to-peak amplitude and zero average is applied to a circuit resembling that in Fig. 4.3(a) and employing a 100-Ω resistor. What is the peak output voltage that results? What is the average output voltage that results? What is the peak diode current? What is the average diode current? What is the maximum reverse voltage across the diode?

4.14 Repeat Problem 4.13 for the situation in which the average voltage of the square wave is 1 V, while its peak-to-peak value remains at 5 V.

D *4.15 Design a battery-charging circuit, resembling that in Fig. 4.4(a) and using an ideal diode, in which current flows to the 12-V battery 25% of the time with an average value of 100 mA. What peak-to-peak sine-wave voltage is required? What resistance is required? What peak diode current flows? What peak reverse voltage does the diode endure? If resistors can be specified to only one significant digit, and the peak-to-peak voltage only to the nearest volt, what design would you choose to guarantee the required charging current? What fraction of the cycle does diode current flow? What is the average diode current? What is the peak diode current? What peak reverse voltage does the diode endure?

4.16 The circuit of Fig. P4.16 can be used in a signaling system using one wire plus a common ground return. At any moment, the input has one of three values: +3 V, 0 V, –3 V.

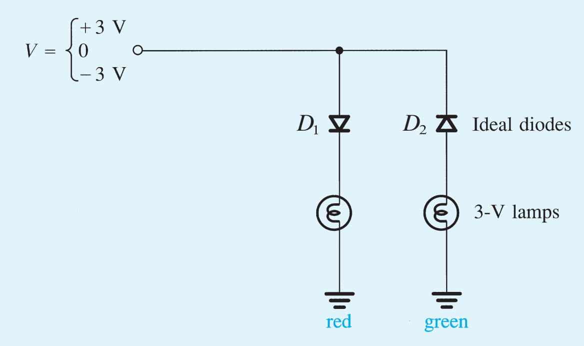


Figure P4.16

What is the status of the lamps for each input value? (Note that the lamps can be located apart from each other and that there may be several of each type of connection, all on one wire!)

Section 4.2: Terminal Characteristics of Junction Diodes

4.17 Calculate the value of the thermal voltage, V_T, at –55°C, 0°C, +40°C, and +125°C. At what temperature is V_T exactly 25 mV?

4.18 At what forward voltage does a diode conduct a current equal to $10{,}000 I_S$? In terms of I_S, what current flows in the same diode when its forward voltage is 0.7 V?

4.19 A diode for which the forward voltage drop is 0.7 V at 1.0 mA is operated at 0.5 V. What is the value of the current?

4.20 A particular diode is found to conduct 1 mA with a junction voltage of 0.7 V. What current will flow in this diode if the junction voltage is raised to 0.71 V? To 0.8 V? If the junction voltage is lowered to 0.69 V? To 0.6 V? What change in junction voltage will increase the diode current by a factor of 10?

4.21 The following measurements are taken on particular junction diodes for which V is the terminal voltage and I is the diode current. For each diode, estimate values of I_S and the terminal voltage at 10% of the measured current.

(a) $V = 0.700$ V at $I = 1.00$ A
(b) $V = 0.650$ V at $I = 1.00$ mA
(c) $V = 0.650$ V at $I = 10$ μA
(d) $V = 0.700$ V at $I = 100$ mA

4.22 Listed below are the results of measurements taken on several different junction diodes. For each diode, the data provided are the diode current I and the corresponding diode voltage V. In each case, estimate I_S, and the diode voltage at $10I$ and $I/10$.

(a) 10.0 mA, 700 mV
(b) 1.0 mA, 700 mV
(c) 10 A, 800 mV
(d) 1 mA, 700 mV
(e) 10 μA, 600 mV

4.23 The circuit in Fig. P4.23 utilizes three identical diodes having $I_S = 10^{-14}$ A. Find the value of the current I required to obtain an output voltage $V_O = 2.0$ V. If a current of 1 mA is drawn away from the output terminal by a load, what is the change in output voltage?

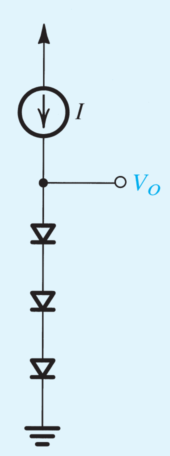


Figure P4.23

4.24 A junction diode is operated in a circuit in which it is supplied with a constant current I. What is the effect on the forward voltage of the diode if an identical diode is connected in parallel?

4.25 Two diodes with saturation currents I_{S1} and I_{S2} are connected in parallel with their cathodes joined together and connected to grounds. The two anodes are joined together and fed with a constant current I. Find the currents I_{D1} and I_{D2} that flow through the two diodes, and the voltage V_D that appears across their parallel combination.

4.26 Four diodes are connected in parallel: anodes joined together and fed with a constant current I, and cathodes joined together and connected to ground. What relative junction areas should these diodes have if their currents must have binary-weighted ratios, with the smallest being 0.1 mA? What value of I is needed?

4.27 In the circuit shown in Fig. P4.27, D_1 has 10 times the junction area of D_2. What value of V results? To obtain a value for V of 60 mV, what current I_2 is needed?

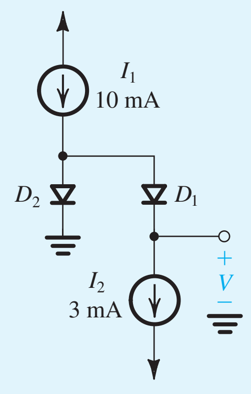


Figure P4.27

4.28 For the circuit shown in Fig. P4.28, both diodes are identical. Find the value of R for which $V = 50$ mV.

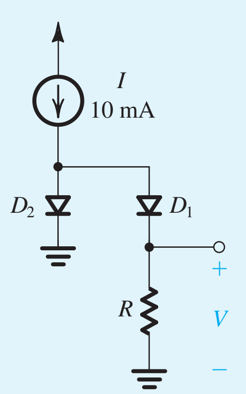


Figure P4.28

4.29 A diode fed with a constant current $I = 1$ mA has a voltage $V = 690$ mV at 20°C. Find the diode voltage at −20°C and at +85°C.

4.30 In the circuit shown in Fig. P4.30, D_1 is a large-area, high-current diode whose reverse leakage is high and independent of applied voltage, while D_2 is a much smaller, low-current diode. At an ambient temperature of 20°C, resistor R_1 is adjusted to make $V_{R1} = V_2 = 520$ mV. Subsequent measurement indicates that R_1 is 520 kΩ. What do you expect the voltages V_{R1} and V_2 to become at 0°C and at 40°C?

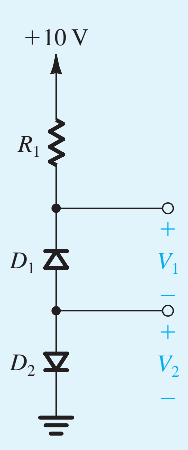


Figure P4.30

4.31 When a 10-A current is applied to a particular diode, it is found that the junction voltage immediately becomes 700 mV. However, as the power being dissipated in the diode raises its temperature, it is found that the voltage decreases and eventually reaches 600 mV. What is the apparent rise in junction temperature? What is the power dissipated in the diode in its final state? What is the temperature rise per watt of power dissipation? (This is called the thermal resistance.)

***4.32** A designer of an instrument that must operate over a wide supply-voltage range, noting that a diode's junction-voltage drop is relatively independent of junction current, considers the use of a large diode to establish a small relatively constant voltage. A power diode, for which the nominal current at 0.8 V is 10 A, is available. If the current source feeding the diode changes in the range 1 mA to 3 mA and if, in addition, the temperature changes by ±20°C, what is the expected range of diode voltage?

***4.33** As an alternative to the idea suggested in Problem 4.32, the designer considers a second approach to producing a relatively constant small voltage from a variable current supply: It relies on the ability to make quite accurate copies of any small current that is available (using a process called current mirroring). The designer proposes to use this idea to supply two diodes of different junction areas with equal currents and to measure their junction-voltage difference. Two types of diodes are available: for a forward voltage of 700 mV, one conducts 0.1 mA, while the other conducts 1 A. Now, for identical currents in the range of 1 mA to 3 mA supplied to each, what range of difference voltages result? What is the effect of a temperature change of ±20°C on this arrangement?

Section 4.3: Modeling the Diode Forward Characteristic

***4.34** Consider the graphical analysis of the diode circuit of Fig. 4.10 with $V_{DD} = 1$ V, $R = 1$ kΩ, and a diode having $I_S = 10^{-15}$ A. Calculate a small number of points on the diode characteristic in the vicinity of where you expect the load line to intersect it, and use a graphical process to refine your estimate of diode current. What value of diode current and voltage do you find? Analytically, find the voltage corresponding to your estimate of current. By how much does it differ from the graphically estimated value?

4.35 Use the iterative-analysis procedure to determine the diode current and voltage in the circuit of Fig. 4.10 for $V_{DD} = 1$ V, $R = 1$ kΩ, and a diode having $I_S = 10^{-15}$ A.

4.36 A "1-mA diode" (i.e., one that has $v_D = 0.7$ V at $i_D = 1$ mA) is connected in series with a 500-Ω resistor to a 1.0 V supply.

(a) Provide a rough estimate of the diode current you would expect.
(b) Estimate the diode current more closely using iterative analysis.

D 4.37 Assuming the availability of diodes for which $v_D = 0.75$ V at $i_D = 1$ mA, design a circuit that utilizes four diodes connected in series, in series with a resistor R connected to a 15-V power supply. The voltage across the string of diodes is to be 3.3 V.

4.38 A diode operates in a series circuit with a resistance R and a dc source V. A designer, considering using a constant-voltage model, is uncertain whether to use 0.7 V or 0.6 V for V_D. For what value of V is the difference in the calculated values of current only 1%? For $V = 3$ V and $R = 1$ kΩ, what two current estimates would result from the use of the two values of V_D? What is their percentage difference?

4.39 A designer has a supply of diodes for which a current of 2 mA flows at 0.7 V. Using a 1-mA current source, the designer wishes to create a reference voltage of 1.3 V. Suggest a combination of series and parallel diodes that will do the job as well as possible. How many diodes are needed? What voltage is actually achieved?

4.40 Solve the problems in Example 4.2 using the constant-voltage-drop ($V_D = 0.7$ V) diode model.

4.41 For the circuits shown in Fig. P4.2, using the constant-voltage-drop ($V_D = 0.7$ V) diode model, find the voltages and currents indicated.

4.42 For the circuits shown in Fig. P4.3, using the constant-voltage-drop ($V_D = 0.7$ V) diode model, find the voltages and currents indicated.

4.43 For the circuits in Fig. P4.9, using the constant-voltage-drop ($V_D = 0.7$ V) diode model, find the values of the labeled currents and voltages.

4.44 For the circuits in Fig. P4.10, utilize Thévenin's theorem to simplify the circuits and find the values of the labeled currents and voltages. Assume that conducting diodes can be represented by the constant-voltage-drop model ($V_D = 0.7$ V).

D 4.45 Repeat Problem 4.11, representing the diode by the constant-voltage-drop ($V_D = 0.7$ V) model. How different is the resulting design?

4.46 The small-signal model is said to be valid for voltage variations of about 5 mV. To what percentage current change does this correspond? (Consider both positive and negative signals.) What is the maximum allowable voltage signal (positive or negative) if the current change is to be limited to 10%?

4.47 In a particular circuit application, ten "20-mA diodes" (a 20-mA diode is a diode that provides a 0.7-V drop when the current through it is 20 mA) connected in parallel operate at a total current of 0.1 A. For the diodes closely matched, what current flows in each? What is the corresponding small-signal resistance of each diode and of the combination? Compare this with the incremental resistance of a single diode conducting 0.1 A. If each of the 20-mA diodes has a series resistance of 0.2 Ω associated with the wire bonds to the junction, what is the equivalent resistance of the 10 parallel-connected diodes? What connection resistance would a single diode need in order to be totally equivalent? (*Note:* This is why the parallel connection of real diodes can often be used to advantage.)

4.48 In the circuit shown in Fig. P4.48, I is a dc current and v_s is a sinusoidal signal. Capacitors C_1 and C_2 are very large; their function is to couple the signal to and from the diode but block the dc current from flowing into the signal source or the load (not shown). Use the diode small-signal model to show that the signal component of the output voltage is

$$v_o = v_s \frac{V_T}{V_T + IR_s}$$

If $v_s = 10$ mV, find v_o for $I = 1$ mA, 0.1 mA, and 1 μA. Let $R_s = 1$ kΩ. At what value of I does v_o become one-half of v_s? Note that this circuit functions as a signal attenuator with the attenuation factor controlled by the value of the dc current I.

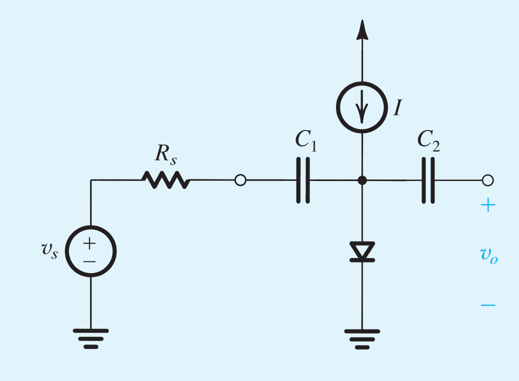


Figure P4.48

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

4.49 In the attenuator circuit of Fig. P4.48, let $R_s = 10\ \text{k}\Omega$. The diode is a 1-mA device; that is, it exhibits a voltage drop of 0.7 V at a dc current of 1 mA. For small input signals, what value of current I is needed for $v_o/v_s = 0.50$? 0.10? 0.01? 0.001? In each case, what is the largest input signal that can be used while ensuring that the signal component of the diode current is limited to ±10% of its dc current? What output signals correspond?

4.50 In the capacitor-coupled attenuator circuit shown in Fig. P4.50, I is a dc current that varies from 0 mA to 1 mA, and C_1 and C_2 are large coupling capacitors. For very small input signals, so that the diodes can be represented by their small-signal resistances r_{d1} and r_{d2}, give the small-signal equivalent circuit and thus show that $\frac{v_o}{v_i} = \frac{r_{d2}}{r_{d1} + r_{d2}}$ and hence that $\frac{v_o}{v_i} = I$, where I is in mA. Find v_o/v_i for $I = 0\ \mu\text{A}$, 1 μA, 10 μA, 100 μA, 500 μA, 600 μA, 900 μA, 990 μA, and 1 mA. Note that this is a signal attenuator whose transmission is linearly controlled by the dc current I.

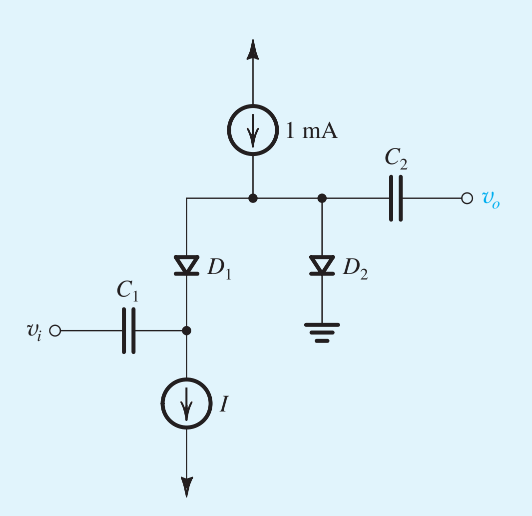


Figure P4.50

***4.51** In the circuit shown in Fig. P4.51, diodes D_1 through D_4 are identical, and each exhibits a voltage drop of 0.7 V at a 1-mA current.

(a) For small input signals (e.g., 10-mV peak), find the small-signal equivalent circuit and use it to determine values of the small-signal transmission v_o/v_i for various values of I: 0 μA, 1 μA, 10 μA, 100 μA, 1 mA, and 10 mA.

(b) For a forward-conducting diode, what is the largest signal-voltage magnitude that it can support while the corresponding signal current is limited to 10% of the dc bias current? Now, for the circuit in Fig. P4.51, for 10-mV peak input, what is the smallest value of I for which the diode currents remain within ±10% of their dc values?

(c) For $I = 1$ mA, what is the largest possible output signal for which the diode currents deviate by at most 10% of their dc values? What is the corresponding peak input? What is the total current in each diode?

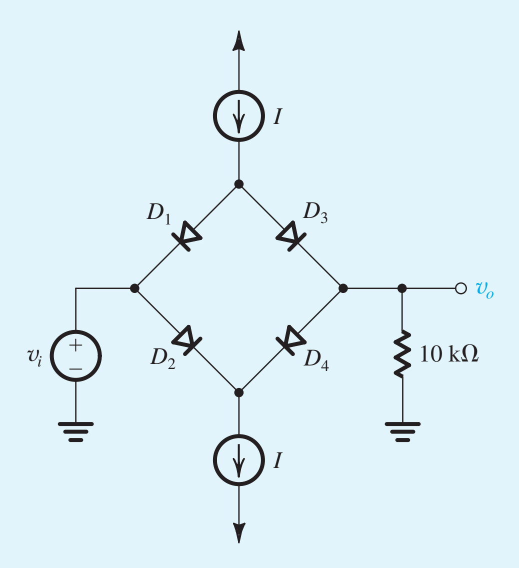


Figure P4.51

****4.52** In Problem 4.51 we investigated the operation of the circuit in Fig. P4.51 for small input signals. In this problem we wish to find the voltage-transfer characteristic (VTC) v_O versus v_I for $-12\text{ V} \le v_I \le 12\text{ V}$ for the case $I = 1$ mA and each of the diodes exhibits a voltage drop of 0.7 V at a current of 1 mA. Toward this end, use the diode exponential characteristic to construct a table that gives the values of: the current i_O in the 10-kΩ resistor, the current in each of the four diodes, the voltage drop across each of the four diodes, and the input voltage v_I, for $v_O = 0$, +1 V, +2 V, +5 V, +9 V, +9.9 V, +9.99 V, +10.5 V, +11 V, and +12 V. Use these data, with extrapolation to negative values of v_I and v_O, to sketch the required VTC. Also sketch the VTC that results if I is reduced to 0.5 mA. (*Hint*: From symmetry, observe that as v_O increases and i_O correspondingly increases, i_{D3} and i_{D2} increase by equal amounts and i_{D4} and i_{D1} decrease by (the same) equal amounts.)

SIM ***4.53** In the circuit shown in Fig. P4.53, I is a dc current and v_i is a sinusoidal signal with small amplitude (less than 10 mV) and a frequency of 100 kHz. Representing the diode by its small-signal resistance r_d, which is a function of I, sketch the small-signal equivalent circuit and use it to determine the sinusoidal output voltage V_o, and thus find the phase shift between V_i and V_o. Find the value of I that will provide a phase shift of –45°, and find the range of phase shift achieved as I is varied over the range of 0.1 times to 10 times this value.

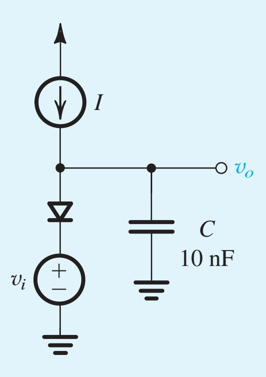


Figure P4.53

***4.54** Consider the voltage-regulator circuit shown in Fig. P4.54. The value of R is selected to obtain an output voltage V_O (across the diode) of 0.7 V.

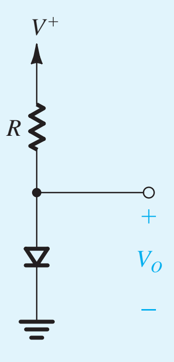


Figure P4.54

(a) Use the diode small-signal model to show that the change in output voltage corresponding to a change of 1 V in V^+ is

$$\frac{\Delta V_O}{\Delta V^+} = \frac{V_T}{V^+ + V_T - 0.7}$$

This quantity is known as the line regulation and is usually expressed in mV/V.

(b) Generalize the expression above for the case of m diodes connected in series and the value of R adjusted so that the voltage across each diode is 0.7 V (and $V_O = 0.7m$ V).

(c) Calculate the value of line regulation for the case $V^+ =$ 15 V (nominally) and (i) $m = 1$ and (ii) $m = 4$.

***4.55** Consider the voltage-regulator circuit shown in Fig P4.54 under the condition that a load current I_L is drawn from the output terminal.

(a) If the value of I_L is sufficiently small that the corresponding change in regulator output voltage ΔV_O is small enough to justify using the diode small-signal model, show that

$$\frac{\Delta V_O}{I_L} = -(r_d \| R)$$

This quantity is known as the load regulation and is usually expressed in mV/mA.

(b) If the value of R is selected such that at no load the voltage across the diode is 0.7 V and the diode current is I_D, show that the expression derived in (a) becomes

$$\frac{\Delta V_O}{I_L} = -\frac{V_T}{I_D}\frac{V^+ - 0.7}{V^+ - 0.7 + V_T}$$

Select the lowest possible value for I_D that results in a load regulation whose magnitude is ≤ 5 mV/mA. If V^+ is nominally 15 V, what value of R is required? Also, specify the diode required in terms of its I_S.

(c) Generalize the expression derived in (b) for the case of m diodes connected in series and R adjusted to obtain $V_O = 0.7m$ V at no load.

D *4.56 Design a diode voltage regulator to supply 1.5 V to a 1.5-kΩ load. Use two diodes specified to have a 0.7-V drop at a current of 1 mA. The diodes are to be connected to a +5-V supply through a resistor R. Specify the value for R. What is the diode current with the load connected? What is the increase resulting in the output voltage when the load is disconnected? What change results if the load resistance is reduced to 1 kΩ? To 750 Ω? To 500 Ω? (*Hint*: Use the small-signal diode model to calculate all changes in output voltage.)

D *4.57 A voltage regulator consisting of two diodes in series fed with a constant-current source is used as a replacement for a single carbon–zinc cell (battery) of nominal voltage 1.5 V. The regulator load current varies from 2 mA to

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

7 mA. Constant-current supplies of 5 mA, 10 mA, and 15 mA are available. Which would you choose, and why? What change in output voltage would result when the load current varies over its full range?

****4.58** A particular design of a voltage regulator is shown in Fig. P4.58. Diodes D_1 and D_2 are 10-mA units; that is, each has a voltage drop of 0.7 V at a current of 10 mA. Use the diode exponential model and iterative analysis to answer the following questions:

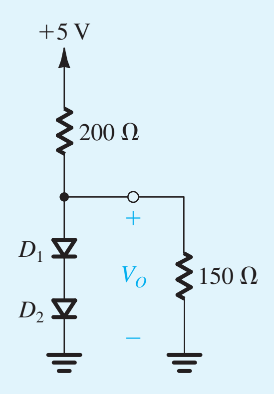


Figure P4.58

(a) What is the regulator output voltage V_O with the 150-Ω load connected?
(b) Find V_O with no load.
(c) With the load connected, to what value can the 5-V supply be lowered while maintaining the loaded output voltage within 0.1 V of its nominal value?
(d) What does the loaded output voltage become when the 5-V supply is raised by the same amount as the drop found in (c)?
(e) For the range of changes explored in (c) and (d), by what percentage does the output voltage change for each percentage change of supply voltage in the worst case?

Section 4.4: Operation in the Reverse Breakdown Region—Zener Diodes

4.59 Partial specifications of a collection of zener diodes are provided below. For each, identify the missing parameter and estimate its value. Note from Fig. 4.19 that $V_{ZK} \simeq V_{Z0}$ and I_{ZK} is very small.

(a) $V_Z = 10.0$ V, $V_{ZK} = 9.6$ V, and $I_{ZT} = 50$ mA
(b) $I_{ZT} = 10$ mA, $V_Z = 9.1$ V, and $r_z = 30\ \Omega$
(c) $r_z = 2\ \Omega$, $V_Z = 6.8$ V, and $V_{ZK} = 6.6$ V
(d) $V_Z = 18$ V, $I_{ZT} = 5$ mA, and $V_{ZK} = 17.6$ V
(e) $I_{ZT} = 200$ mA, $V_Z = 7.5$ V, and $r_z = 1.5\ \Omega$

Assuming that the power rating of a breakdown diode is established at about twice the specified zener current (I_{ZT}), what is the power rating of each of the diodes described above?

D 4.60 A designer requires a shunt regulator of approximately 20 V. Two kinds of zener diodes are available: 6.8-V devices with r_z of 10 Ω and 5.1-V devices with r_z of 25 Ω. For the two major choices possible, find the load regulation. In this calculation neglect the effect of the regulator resistance R.

4.61 A shunt regulator utilizing a zener diode with an incremental resistance of 8 Ω is fed through an 82-Ω resistor. If the raw supply changes by 1.0 V, what is the corresponding change in the regulated output voltage?

4.62 A 9.1-V zener diode exhibits its nominal voltage at a test current of 20 mA. At this current the incremental resistance is specified as 10 Ω. Find V_{Z0} of the zener model. Find the zener voltage at a current of 10 mA and at 50 mA.

D 4.63 Design a 7.5-V zener regulator circuit using a 7.5-V zener specified at 10 mA. The zener has an incremental resistance $r_z = 30\ \Omega$ and a knee current of 0.5 mA. The regulator operates from a 10-V supply and has a 1.5-kΩ load. What is the value of R you have chosen? What is the regulator output voltage when the supply is 10% high? Is 10% low? What is the output voltage when both the supply is 10% high and the load is removed? What is the smallest possible load resistor that can be used while the zener operates at a current no lower than the knee current while the supply is 10% low? What is the load voltage in this case?

D 4.64 Provide two designs of shunt regulators utilizing the 1N5235 zener diode, which is specified as follows: $V_Z = 6.8$ V and $r_z = 5\ \Omega$ for $I_Z = 20$ mA; at $I_Z = 0.25$ mA (nearer the knee), $r_z = 750\ \Omega$. For both designs, the supply voltage is nominally 9 V and varies by ±1 V. For the first design, assume that the availability of supply current is not a problem, and thus operate the diode at 20 mA. For the second design, assume that the current from the raw supply is limited, and therefore you are forced to operate the diode at 0.25 mA. For the purpose of these initial designs, assume no load. For each design find the value of R and the line regulation.

D *4.65 A zener shunt regulator employs a 9.1-V zener diode for which $V_Z = 9.1$ V at $I_Z = 9$ mA, with $r_z = 40\ \Omega$ and

$I_{ZK} = 0.5$ mA. The available supply voltage of 15 V can vary as much as ±10%. For this diode, what is the value of V_{Z0}? For a nominal load resistance R_L of 1 kΩ and a nominal zener current of 10 mA, what current must flow in the supply resistor R? For the nominal value of supply voltage, select a value for resistor R, specified to one significant digit, to provide at least that current. What nominal output voltage results? For a ±10% change in the supply voltage, what variation in output voltage results? If the load current is reduced by 50%, what increase in V_O results? What is the smallest value of load resistance that can be tolerated while maintaining regulation when the supply voltage is low? What is the lowest possible output voltage that results? Calculate values for the line regulation and for the load regulation for this circuit using the numerical results obtained in this problem.

D *4.66 It is required to design a zener shunt regulator to provide a regulated voltage of about 10 V. The available 10-V, 1-W zener of type 1N4740 is specified to have a 10-V drop at a test current of 25 mA. At this current, its r_z is 7 Ω. The raw supply, V_S, available has a nominal value of 20 V but can vary by as much as ±25%. The regulator is required to supply a load current of 0 mA to 20 mA. Design for a minimum zener current of 5 mA.

(a) Find V_{Z0}.
(b) Calculate the required value of R.
(c) Find the line regulation. What is the change in V_O expressed as a percentage, corresponding to the ±25% change in V_S?
(d) Find the load regulation. By what percentage does V_O change from the no-load to the full-load condition?
(e) What is the maximum current that the zener in your design is required to conduct? What is the zener power dissipation under this condition?

Section 4.5: Rectifier Circuits

4.67 Consider the half-wave rectifier circuit of Fig. 4.23(a) with the diode reversed. Let v_S be a sinusoid with 10-V peak amplitude, and let $R = 1$ kΩ. Use the constant-voltage-drop diode model with $V_D = 0.7$ V.

(a) Sketch the transfer characteristic.
(b) Sketch the waveform of v_O.
(c) Find the average value of v_O.
(d) Find the peak current in the diode.
(e) Find the PIV of the diode.

4.68 Using the exponential diode characteristic, show that for v_S and v_O both greater than zero, the circuit of Fig. 4.23(a) has the transfer characteristic

$$v_O = v_S - v_D\left(\text{at } i_D = 1 \text{ mA}\right) - V_T \ln\left(v_O/R\right)$$

where v_S and v_O are in volts and R is in kilohms. Note that this relationship can be used to obtain the voltage transfer characteristic v_O vs. v_S by finding v_S corresponding to various values of v_O.

SIM **4.69** Consider a half-wave rectifier circuit with a triangular-wave input of 5-V peak-to-peak amplitude and zero average, and with $R = 1$ kΩ. Assume that the diode can be represented by the constant-voltage-drop model with $V_D = 0.7$ V. Find the average value of v_O.

4.70 A half-wave rectifier circuit with a 1-kΩ load operates from a 120-V (rms) 60-Hz household supply through a 12-to-1 step-down transformer. It uses a silicon diode that can be modeled to have a 0.7-V drop for any current. What is the peak voltage of the rectified output? For what fraction of the cycle does the diode conduct? What is the average output voltage? What is the average current in the load?

4.71 A full-wave rectifier circuit with a 1-kΩ load operates from a 120-V (rms) 60-Hz household supply through a 6-to-1 transformer having a center-tapped secondary winding. It uses two silicon diodes that can be modeled to have a 0.7-V drop for all currents. What is the peak voltage of the rectified output? For what fraction of a cycle does each diode conduct? What is the average output voltage? What is the average current in the load?

4.72 A full-wave bridge-rectifier circuit with a 1-kΩ load operates from a 120-V (rms) 60-Hz household supply through a 12-to-1 step-down transformer having a single secondary winding. It uses four diodes, each of which can be modeled to have a 0.7-V drop for any current. What is the peak value of the rectified voltage across the load? For what fraction of a cycle does each diode conduct? What is the average voltage across the load? What is the average current through the load?

D 4.73 It is required to design a full-wave rectifier circuit using the circuit of Fig. 4.24 to provide an average output voltage of:

(a) 10 V
(b) 100 V

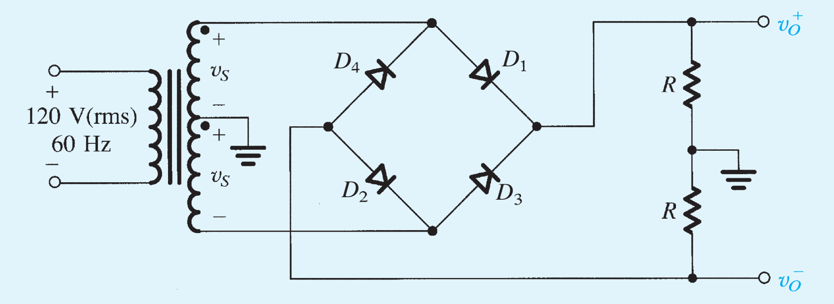


Figure P4.76

In each case find the required turns ratio of the transformer. Assume that a conducting diode has a voltage drop of 0.7 V. The ac line voltage is 120 V rms.

D 4.74 Repeat Problem 4.73 for the bridge-rectifier circuit of Fig. 4.25.

D 4.75 Consider the full-wave rectifier in Fig. 4.24 when the transformer turns ratio is such that the voltage across the entire secondary winding is 20 V rms. If the input ac line voltage (120 V rms) fluctuates by as much as ±10%, find the required PIV of the diodes. (Remember to use a factor of safety in your design.)

4.76 The circuit in Fig. P4.76 implements a complementary-output rectifier. Sketch and clearly label the waveforms of v_O^+ and v_O^-. Assume a 0.7-V drop across each conducting diode. If the magnitude of the average of each output is to be 12 V, find the required amplitude of the sine wave across the entire secondary winding. What is the PIV of each diode?

4.77 Augment the rectifier circuit of Problem 4.70 with a capacitor chosen to provide a peak-to-peak ripple voltage of (i) 10% of the peak output and (ii) 1% of the peak output. In each case:

(a) What average output voltage results?
(b) What fraction of the cycle does the diode conduct?
(c) What is the average diode current?
(d) What is the peak diode current?

4.78 Repeat Problem 4.77 for the rectifier in Problem 4.71.

4.79 Repeat Problem 4.77 for the rectifier in Problem 4.72.

D *4.80 It is required to use a peak rectifier to design a dc power supply that provides an average dc output voltage of 12 V on which a maximum of ±1-V ripple is allowed. The rectifier feeds a load of 200 Ω. The rectifier is fed from the line voltage (120 V rms, 60 Hz) through a transformer. The diodes available have 0.7-V drop when conducting. If the designer opts for the half-wave circuit:

(a) Specify the rms voltage that must appear across the transformer secondary.
(b) Find the required value of the filter capacitor.
(c) Find the maximum reverse voltage that will appear across the diode, and specify the PIV rating of the diode.
(d) Calculate the average current through the diode during conduction.
(e) Calculate the peak diode current.

D *4.81 Repeat Problem 4.80 for the case in which the designer opts for a full-wave circuit utilizing a center-tapped transformer.

D *4.82 Repeat Problem 4.80 for the case in which the designer opts for a full-wave bridge-rectifier circuit.

D *4.83 Consider a half-wave peak rectifier fed with a voltage v_S having a triangular waveform with 24-V peak-to-peak amplitude, zero average, and 1-kHz frequency. Assume that the diode has a 0.7-V drop when conducting. Let the load resistance $R = 100\ \Omega$ and the filter capacitor $C = 100\ \mu F$. Find the average dc output voltage, the time interval during which the diode conducts, the average diode current during conduction, and the maximum diode current.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

D *4.84 Consider the circuit in Fig. P4.76 with two equal filter capacitors placed across the load resistors R. Assume that the diodes available exhibit a 0.7-V drop when conducting. Design the circuit to provide ±12-V dc output voltages with a peak-to-peak ripple no greater than 1 V. Each supply should be capable of providing 100-mA dc current to its load resistor R. Completely specify the capacitors, diodes, and the transformer.

4.85 The op amp in the precision rectifier circuit of Fig. P4.85 is ideal with output saturation levels of ±13 V. Assume that when conducting the diode exhibits a constant voltage drop of 0.7 V. Find v_-, v_O, and v_A for:

(a) $v_I = +1$ V
(b) $v_I = +3$ V
(c) $v_I = -1$ V
(d) $v_I = -3$ V

Also, find the average output voltage obtained when v_I is a symmetrical square wave of 1-kHz frequency, 5-V amplitude, and zero average.

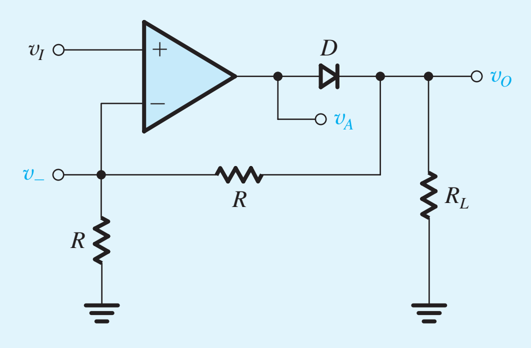


Figure P4.85

4.86 The op amp in the circuit of Fig. P4.86 is ideal with output saturation levels of ±12 V. The diodes exhibit a constant 0.7-V drop when conducting. Find v_-, v_A, and v_O for:

(a) $v_I = +1$ V
(b) $v_I = +3$ V
(c) $v_I = -1$ V
(d) $v_I = -3$ V

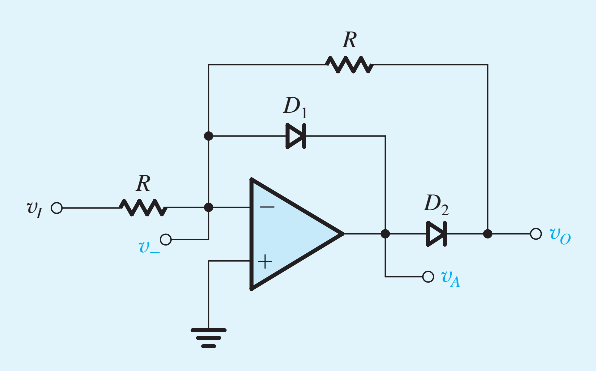
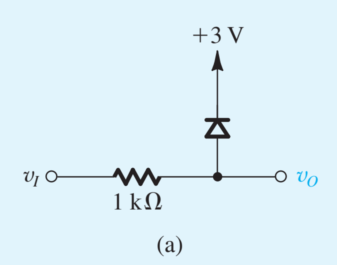


Figure P4.86

Section 4.6: Limiting and Clamping Circuits

4.87 Sketch the transfer characteristic v_O versus v_I for the limiter circuits shown in Fig. P4.87. All diodes begin conducting at a forward voltage drop of 0.5 V and have voltage drops of 0.7 V when conducting a current $i_D \geq 1$ mA.



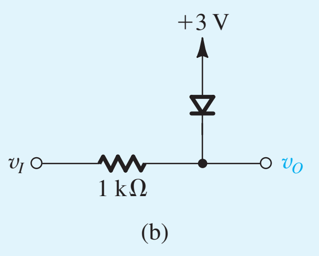


Figure P4.87

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

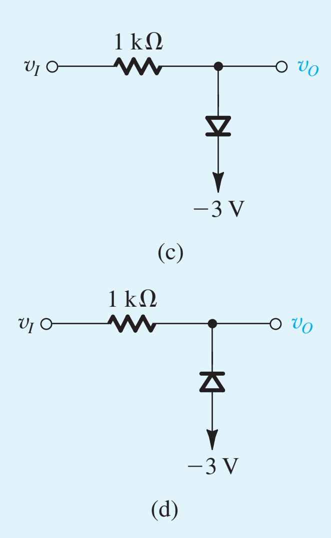


Figure P4.87 *continued*

4.88 The circuits in Fig. P4.87(a) and (d) are connected as follows: The two input terminals are tied together, and the two output terminals are tied together. Sketch the transfer characteristic of the circuit resulting, assuming that the cut-in voltage of the diodes is 0.5 V and their voltage drop when conducting a current $i_D \geq 1$ mA is 0.7 V.

4.89 Repeat Problem 4.88 for the two circuits in Fig. P4.87(a) and (b) connected together as follows: The two input terminals are tied together, and the two output terminals are tied together.

4.90 Sketch and clearly label the transfer characteristic of the circuit in Fig. P4.90 for $-15\text{ V} \leq v_I \leq +15$ V. Assume that the diodes can be represented by the constant-voltage-drop model with $V_D = 0.7$ V. Also assume that the zener voltage is 6.8 V and that r_z is negligibly small.

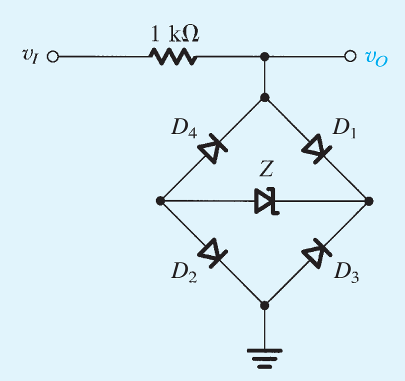


Figure P4.90

***4.91** Plot the transfer characteristic of the circuit in Fig. P4.91 by evaluating v_I corresponding to $v_O = 0.5$ V, 0.6 V, 0.7 V, 0.8 V, 0 V, –0.5 V, –0.6 V, –0.7 V, and –0.8 V. Use the exponential model for the diodes, and assume that they have 0.7-V drops at 1-mA currents. Characterize the circuit as a hard or soft limiter. What is the value of K? Estimate L_+ and L_-.

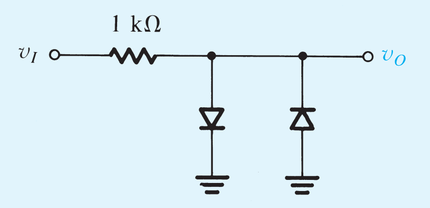


Figure P4.91

4.92 Design limiter circuits using only diodes and 10-kΩ resistors to provide an output signal limited to the range:

(a) –0.7 V and above
(b) +2.1 V and below
(c) ±1.4 V

Assume that each diode has a 0.7-V drop when conducting.

4.93 Design a two-sided limiting circuit using a resistor, two diodes, and two power supplies to feed a 1-kΩ load with nominal limiting levels of ±2.2 V. Use diodes modeled by a constant 0.7 V. In the nonlimiting region, the voltage gain should be at least 0.94 V/V.

****4.94** In the circuit shown in Fig. P4.94, the diodes exhibit a 0.7-V drop at 0.1 mA. For inputs over the range of ±5 V, use the diode exponential model to provide a calibrated sketch of the voltages at outputs B and C versus v_A. For a 5-V peak,

100-Hz sinusoid applied at A, sketch the signals at nodes B and C.

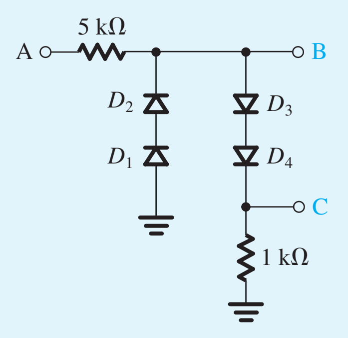


Figure P4.94

**4.95 Sketch and label the voltage-transfer characteristic v_O versus v_I of the circuit shown in Fig. P4.95 over a ±10-V range of input signals. Use the diode exponential model and assume that all diodes are 1-mA units (i.e., each exhibits a 0.7-V drop at a current of 1 mA). What are the slopes of the characteristics at the extreme ±10-V levels?

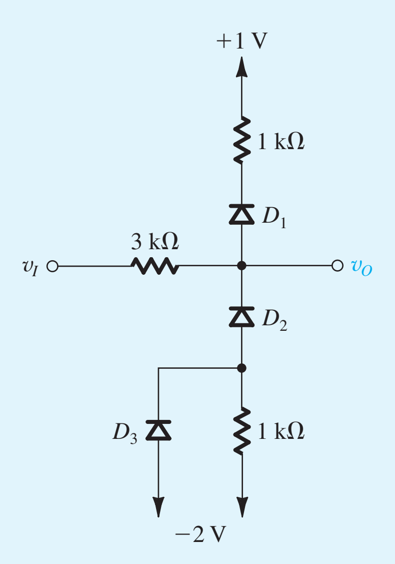


Figure P4.95

4.96 A clamped capacitor using an ideal diode with cathode grounded is supplied with a sine wave of 5-V rms. What is the average (dc) value of the resulting output?

*4.97 For the circuits in Fig. P4.97, each utilizing an ideal diode (or diodes), sketch the output for the input shown. Label the most positive and most negative output levels. Assume $CR \gg T$.

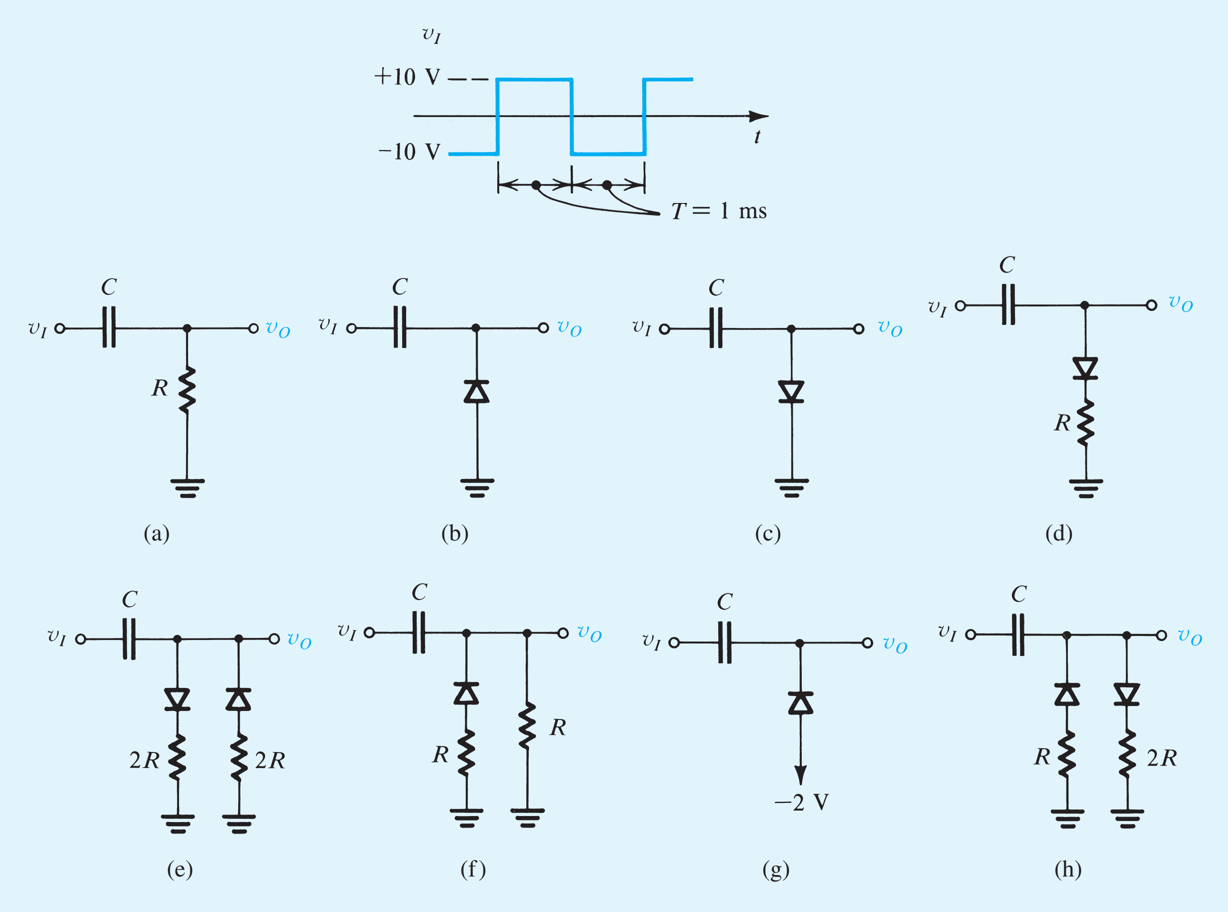


Figure P4.97

CHAPTER 5

MOS Field-Effect Transistors (MOSFETs)

IN THIS CHAPTER YOU WILL LEARN

1. The physical structure of the MOS transistor and how it works.
2. How the voltage between two terminals of the transistor controls the current that flows through the third terminal, and the equations that describe these current–voltage characteristics.
3. How to analyze and design circuits that contain MOS transistors, resistors, and dc sources.

Introduction

Having studied the junction diode, which is the most basic two-terminal semiconductor device, we now turn our attention to three-terminal semiconductor devices. Three-terminal devices are far more useful than two-terminal ones because they can be used in a multitude of applications, ranging from signal amplification to digital logic and memory. The basic principle involved is the use of the voltage between two terminals to control the current flowing in the third terminal. In this way a three-terminal device can be used to realize a controlled source, which as we have learned in Chapter 1 is the basis for amplifier design. Also, in the extreme, the control signal can be used to cause the current in the third terminal to change from zero to a large value, thus allowing the device to act as a switch. As we shall see in Chapter 14, the switch is the basis for the realization of the logic inverter, the basic element of digital circuits.

There are two major types of three-terminal semiconductor devices: the metal-oxide-semiconductor field-effect transistor (MOSFET), which is studied in this chapter, and the bipolar junction transistor (BJT), which we shall study in Chapter 6. Although each of the two transistor types offers unique features and areas of application, the MOSFET has become by far the most widely used electronic device, especially in the design of integrated circuits (ICs), which are entire circuits fabricated on a single silicon chip.

Compared to BJTs, MOSFETs can be made quite small (i.e., requiring a small area on the silicon IC chip), and their manufacturing process is relatively simple (see Appendix A). Also, their operation requires comparatively little power. Furthermore, circuit designers have found ingenious ways to implement digital and analog functions utilizing MOSFETs almost exclusively (i.e., with very few or no resistors). All of these properties have made it possible to pack large numbers of MOSFETs (as many as 4 billion!) on a single IC chip to implement very sophisticated, very-large-scale-integrated (VLSI) digital circuits such as those for memory and microprocessors. Analog circuits such as amplifiers and filters can also be implemented in MOS technology, albeit in smaller, less-dense chips. Also, both analog and digital functions are increasingly being implemented on the same IC chip, in what is known as mixed-signal design.

The objective of this chapter is to develop in the reader a high degree of familiarity with the MOSFET: its physical structure and operation, terminal characteristics, and dc circuit applications. This will provide a solid foundation for the application of the MOSFET in amplifier design (Chapter 7) and in digital circuit design (Chapter 14). Although discrete MOS transistors exist, and the material studied in this chapter will enable the reader to design discrete MOS circuits, our study of the MOSFET is strongly influenced by the fact that most of its applications are in integrated-circuit design. The design of IC analog and digital MOS circuits occupies a large proportion of the remainder of this book.

THE FIRST FIELD-EFFECT DEVICES:

In 1925 a patent for solid-state electric-field-controlled conductor was filed in Canada by Julius E. Lilienfeld, a physicist at the University of Leipzig, Germany. Other patent refinements followed in the United States in 1926 and 1928. Regrettably, no research papers were published. Consequently, in 1934 Oskar Heil, a German physicist working at the University of Cambridge, U.K., filed a patent on a similar idea. But all these early concepts of electric-field control of a semiconducting path languished because suitable technology was not available.

The invention of the bipolar transistor in 1947 at Bell Telephone Laboratories resulted in the speedy development of bipolar devices, a circumstance that further delayed the development of field-effect transistors. Although the field-effect device was described in a paper by William Shockley in 1952, it was not until 1960 that a patent on an insulated-gate field-effect device, the MOSFET, was filed by Dawon Kahng and Martin Atalla, also at Bell Labs. Clearly, the idea of field-effect control for amplification and switching has changed the world. With integrated-circuit chips today containing billions of MOS devices, MOS dominates the electronics world!

5.1 Device Structure and Physical Operation

The enhancement-type MOSFET is the most widely used field-effect transistor. Except for the last section, this chapter is devoted to the study of the enhancement-type MOSFET. We begin in this section by learning about its structure and physical operation. This will lead to the current–voltage characteristics of the device, studied in the next section.

5.1.1 Device Structure

Figure 5.1 shows the physical structure of the *n*-channel enhancement-type MOSFET. The meaning of the names "enhancement" and "*n*-channel" will become apparent shortly. The transistor is fabricated on a *p*-type substrate, which is a single-crystal silicon wafer that provides physical support for the device (and for the entire circuit in the case of an integrated circuit). Two heavily doped *n*-type regions, indicated in the figure as the n^+ **source**[1] and the n^+ **drain** regions, are created in the substrate. A thin layer of silicon dioxide (SiO_2) of thickness t_{ox} (typically 1 nm to 10 nm),[2] which is an excellent electrical insulator, is grown on the surface of the substrate, covering the area between the source and drain regions. Metal is deposited on top of the oxide layer to form the **gate electrode** of the device. Metal contacts are

[1]The notation n^+ indicates heavily doped *n*-type silicon. Conversely, n^- is used to denote lightly doped *n*-type silicon. Similar notation applies for *p*-type silicon.

[2]A nanometer (nm) is 10^{-9} m or 0.001 μm. A micrometer (μm), or micron, is 10^{-6} m. Sometimes the oxide thickness is expressed in angstroms. An angstrom (Å) is 10^{-1} nm, or 10^{-10} m.

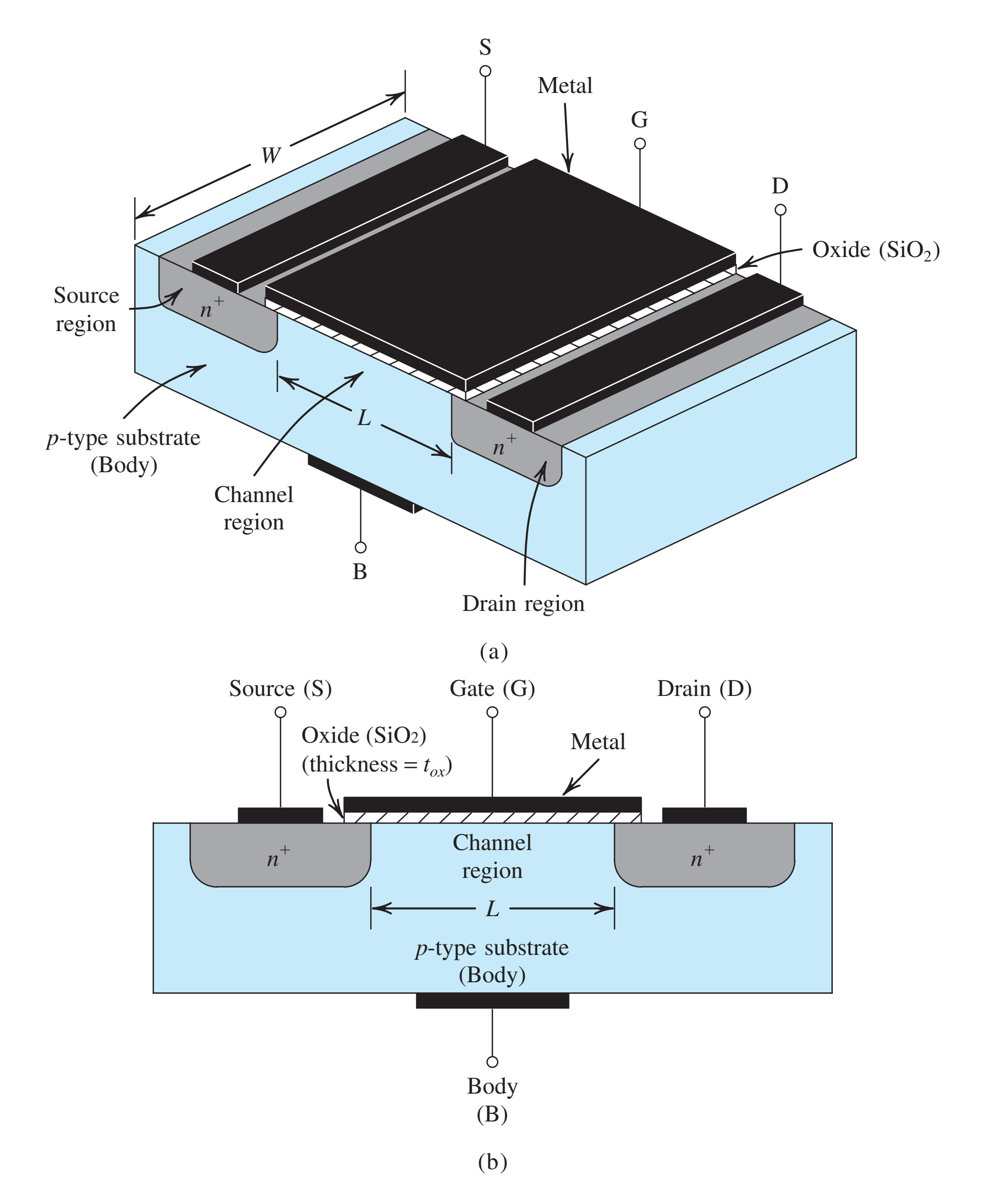


Figure 5.1 Physical structure of the enhancement-type NMOS transistor: **(a)** perspective view; **(b)** cross section. Typically $L = 0.03\ \mu m$ to $1\ \mu m$, $W = 0.05\ \mu m$ to $100\ \mu m$, and the thickness of the oxide layer (t_{ox}) is in the range of 1 to 10 nm.

also made to the source region, the drain region, and the substrate, also known as the **body**.[3] Thus four terminals are brought out: the gate terminal (G), the source terminal (S), the drain terminal (D), and the substrate or body terminal (B).

[3]In Fig. 5.1, the contact to the body is shown on the bottom of the device. This will prove helpful in Section 5.4 in explaining a phenomenon known as the "body effect." It is important to note, however, that in actual ICs, contact to the body is made at a location on the top of the device.

At this point it should be clear that the name of the device (metal-oxide-semiconductor FET) is derived from its physical structure. The name, however, has become a general one and is used also for FETs that do not use metal for the gate electrode. In fact, most modern MOSFETs are fabricated using a process known as silicon-gate technology, in which a certain type of silicon, called polysilicon, is used to form the gate electrode (see Appendix A). Our description of MOSFET operation and characteristics applies irrespective of the type of gate electrode.

Another name for the MOSFET is the **insulated-gate FET** or **IGFET**. This name also arises from the physical structure of the device, emphasizing the fact that the gate electrode is electrically insulated from the device body (by the oxide layer). It is this insulation that causes the current in the gate terminal to be extremely small (of the order of 10^{-15} A).

Observe that the substrate forms *pn* junctions with the source and drain regions. In normal operation these *pn* junctions are kept reverse biased at all times. Since, as we shall see shortly, the drain will always be at a positive voltage relative to the source, the two *pn* junctions can be effectively cut off by simply connecting the substrate terminal to the source terminal. We shall assume this to be the case in the following description of MOSFET operation. Thus, here, the substrate will be considered as having no effect on device operation, and the MOSFET will be treated as a three-terminal device, with the terminals being the gate (G), the source (S), and the drain (D). It will be shown that a voltage applied to the gate controls current flow between source and drain. This current will flow in the longitudinal direction from drain to source in the region labeled "channel region." Note that this region has a length L and a width W, two important parameters of the MOSFET. Typically, L is in the range of 0.03 μm to 1 μm, and W is in the range of 0.05 μm to 100 μm. Finally, note that the MOSFET is a symmetrical device; thus its source and drain can be interchanged with no change in device characteristics.

5.1.2 Operation with Zero Gate Voltage

With zero voltage applied to the gate, two back-to-back diodes exist in series between drain and source. One diode is formed by the *pn* junction between the n^+ drain region and the *p*-type substrate, and the other diode is formed by the *pn* junction between the *p*-type substrate and the n^+ source region. These back-to-back diodes prevent current conduction from drain to source when a voltage v_{DS} is applied. In fact, the path between drain and source has a very high resistance (of the order of 10^{12} Ω).

5.1.3 Creating a Channel for Current Flow

Consider next the situation depicted in Fig. 5.2. Here we have grounded the source and the drain and applied a positive voltage to the gate. Since the source is grounded, the gate voltage appears in effect between gate and source and thus is denoted v_{GS}. The positive voltage on the gate causes, in the first instance, the free holes (which are positively charged) to be repelled from the region of the substrate under the gate (the channel region). These holes are pushed downward into the substrate, leaving behind a carrier-depletion region. The depletion region is populated by the bound negative charge associated with the acceptor atoms. These charges are "uncovered" because the neutralizing holes have been pushed downward into the substrate.

As well, the positive gate voltage attracts electrons from the n^+ source and drain regions (where they are in abundance) into the channel region. When a sufficient number of electrons accumulate near the surface of the substrate under the gate, an n region is in effect created, connecting the source and drain regions, as indicated in Fig. 5.2. Now if a voltage is applied between drain and source, current flows through this induced n region, carried by the mobile

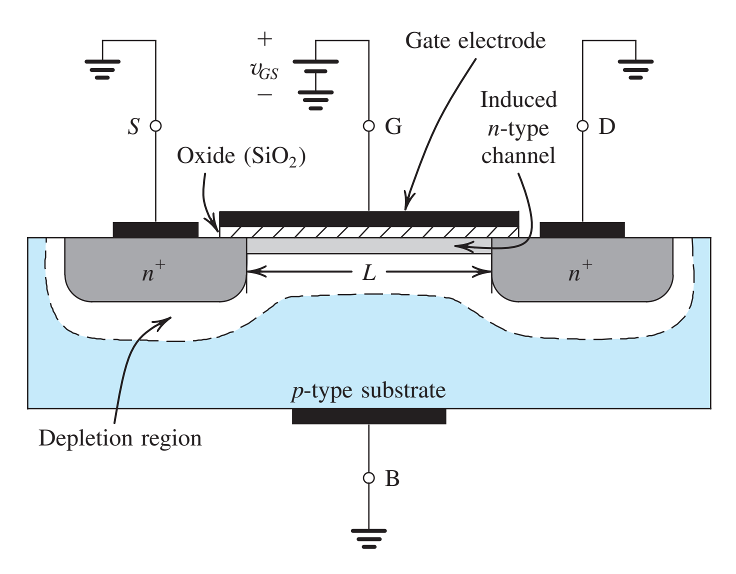


Figure 5.2 The enhancement-type NMOS transistor with a positive voltage applied to the gate. An *n* channel is induced at the top of the substrate beneath the gate.

electrons. The *induced n* region thus forms a **channel** for current flow from drain to source and is aptly called so. Correspondingly, the MOSFET of Fig. 5.2 is called an ***n*-channel MOSFET** or, alternatively, an **NMOS transistor**. Note that an *n*-channel MOSFET is formed in a *p*-type substrate: The channel is created by *inverting* the substrate surface from *p* type to *n* type. Hence the induced channel is also called an **inversion layer**.

The value of v_{GS} at which a sufficient number of mobile electrons accumulate in the channel region to form a conducting channel is called the **threshold voltage** and is denoted V_t.[4] Obviously, V_t for an *n*-channel FET is positive. The value of V_t is controlled during device fabrication and typically lies in the range of 0.3 V to 1.0 V.

The gate and the channel region of the MOSFET form a parallel-plate capacitor, with the oxide layer acting as the capacitor dielectric. The positive gate voltage causes positive charge to accumulate on the top plate of the capacitor (the gate electrode). The corresponding negative charge on the bottom plate is formed by the electrons in the induced channel. An electric field thus develops in the vertical direction. It is this field that controls the amount of charge in the channel, and thus it determines the channel conductivity and, in turn, the current that will flow through the channel when a voltage v_{DS} is applied. This is the origin of the name "field-effect transistor" (FET).

The voltage across this parallel-plate capacitor, that is, the voltage across the oxide, must exceed V_t for a channel to form. When $v_{DS} = 0$, as in Fig. 5.2, the voltage at every point along the channel is zero, and the voltage across the oxide (i.e., between the gate and the points along the channel) is uniform and equal to v_{GS}. The excess of v_{GS} over V_t is termed the **effective voltage** or the **overdrive voltage** and is the quantity that determines the charge in the channel. In this book, we shall denote $(v_{GS} - V_t)$ by v_{OV},

$$v_{GS} - V_t \equiv v_{OV} \tag{5.1}$$

[4]Some texts use V_T to denote the threshold voltage. We use V_t to avoid confusion with the thermal voltage V_T.

We can express the magnitude of the electron charge in the channel by

$$|Q| = C_{ox}(WL)v_{OV} \tag{5.2}$$

where C_{ox}, called the **oxide capacitance**, is the capacitance of the parallel-plate capacitor per unit gate area (in units of F/m^2), W is the width of the channel, and L is the length of the channel. The oxide capacitance C_{ox} is given by

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} \tag{5.3}$$

where ϵ_{ox} is the permittivity of the silicon dioxide,

$$\epsilon_{ox} = 3.9\epsilon_0 = 3.9 \times 8.854 \times 10^{-12} = 3.45 \times 10^{-11} \text{ F/m}$$

The oxide thickness t_{ox} is determined by the process technology used to fabricate the MOSFET. As an example, for a process with $t_{ox} = 4$ nm,

$$C_{ox} = \frac{3.45 \times 10^{-11}}{4 \times 10^{-9}} = 8.6 \times 10^{-3} \text{ F/m}^2$$

It is much more convenient to express C_{ox} per micron squared. For our example, this yields 8.6 fF/μm^2, where fF denotes femtofarad (10^{-15} F). For a MOSFET fabricated in this technology with a channel length $L = 0.18$ μm and a channel width $W = 0.72$ μm, the total capacitance between gate and channel is

$$C = C_{ox}WL = 8.6 \times 0.18 \times 0.72 = 1.1 \text{ fF}$$

Finally, note from Eq. (5.2) that as v_{OV} is increased, the magnitude of the channel charge increases proportionately. Sometimes this is depicted as an increase in the depth of the channel; that is, the larger the overdrive voltage, the deeper the channel.

5.1.4 Applying a Small v_{DS}

Having induced a channel, we now apply a positive voltage v_{DS} between drain and source, as shown in Fig. 5.3. We first consider the case where v_{DS} is small (i.e., 50 mV or so). The voltage v_{DS} causes a current i_D to flow through the induced n channel. Current is carried by free electrons traveling from source to drain (hence the names source and drain). By convention, the direction of current flow is opposite to that of the flow of negative charge. Thus the current in the channel, i_D, will be from drain to source, as indicated in Fig. 5.3.

We now wish to calculate the value of i_D. Toward that end, we first note that because v_{DS} is small, we can continue to assume that the voltage between the gate and various points along the channel remains approximately constant and equal to the value at the source end, v_{GS}. Thus, the effective voltage between the gate and the various points along the channel remains equal to v_{OV}, and the channel charge Q is still given by Eq. (5.2). Of particular interest

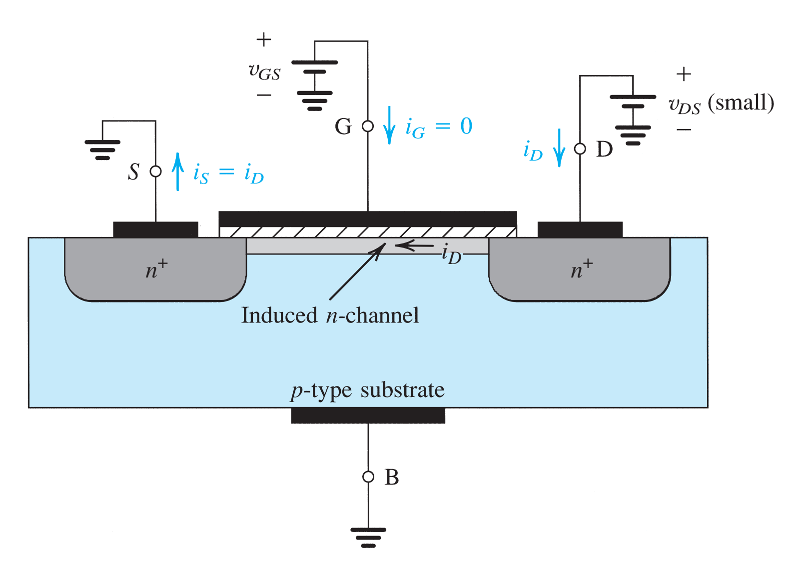


Figure 5.3 An NMOS transistor with $v_{GS} > V_t$ and with a small v_{DS} applied. The device acts as a resistance whose value is determined by v_{GS}. Specifically, the channel conductance is proportional to $v_{GS} - V_t$, and thus i_D is proportional to $(v_{GS} - V_t)v_{DS}$. Note that the depletion region is not shown (for simplicity).

in calculating the current i_D is the charge per unit channel length, which can be found from Eq. (5.2) as

$$\frac{|Q|}{\text{unit channel length}} = C_{ox} W v_{OV} \tag{5.4}$$

The voltage v_{DS} establishes an electric field E across the length of the channel,

$$|E| = \frac{v_{DS}}{L} \tag{5.5}$$

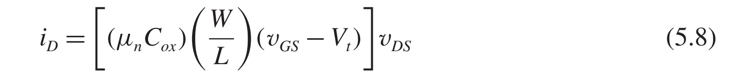
This electric field in turn causes the channel electrons to drift toward the drain with a velocity given by

$$\text{Electron drift velocity} = \mu_n |E| = \mu_n \frac{v_{DS}}{L} \tag{5.6}$$

where μ_n is the mobility of the electrons at the surface of the channel. It is a physical parameter whose value depends on the fabrication process technology. The value of i_D can now be found by multiplying the charge per unit channel length (Eq. 5.4) by the electron drift velocity (Eq. 5.6),

$$i_D = \left[(\mu_n C_{ox})\left(\frac{W}{L}\right) v_{OV}\right] v_{DS} \tag{5.7}$$

Thus, for small v_{DS}, the channel behaves as a linear resistance whose value is controlled by the overdrive voltage v_{OV}, which in turn is determined by v_{GS}:



$$i_D = \left[(\mu_n C_{ox})\left(\frac{W}{L}\right)(v_{GS} - V_t)\right] v_{DS} \tag{5.8}$$

The conductance g_{DS} of the channel can be found from Eq. (5.7) or (5.8) as

$$g_{DS} = (\mu_n C_{ox})\left(\frac{W}{L}\right) v_{OV} \tag{5.9}$$

or

$$g_{DS} = (\mu_n C_{ox})\left(\frac{W}{L}\right)(v_{GS} - V_t) \tag{5.10}$$

Observe that the conductance is determined by the product of three factors: $(\mu_n C_{ox})$, (W/L), and v_{OV} (or equivalently, $v_{GS} - V_t$). To gain insight into MOSFET operation, we consider each of the three factors in turn.

The first factor, $(\mu_n C_{ox})$, is determined by the process technology used to fabricate the MOSFET. It is the product of the electron mobility, μ_n, and the oxide capacitance, C_{ox}. It makes physical sense for the channel conductance to be proportional to each of μ_n and C_{ox} (why?) and hence to their product, which is termed the **process transconductance** parameter[5] and given the symbol k_n', where the subscript n denotes n channel,

$$k_n' = \mu_n C_{ox} \tag{5.11}$$

It can be shown that with μ_n having the dimensions of meters squared per volt-second ($m^2/V \cdot s$) and C_{ox} having the dimensions of farads per meter squared (F/m^2), the dimensions of k_n' are amperes per volt squared (A/V^2).

The second factor in the expression for the conductance g_{DS} in Eqs. (5.9) and (5.10) is the transistor **aspect ratio** (W/L). That the channel conductance is proportional to the channel width W and inversely proportional to the channel length L should make perfect physical sense. The (W/L) ratio is obviously a dimensionless quantity that is determined by the device designer. Indeed, the values of W and L can be selected by the device designer to give the device the i–v characteristics desired. For a given fabrication process, however, there is a minimum channel length, L_{min}. In fact, the minimum channel length that is possible with a given fabrication process is used to characterize the process and is being continually reduced as technology advances. For instance, in 2014 the state-of-the-art in commercially available MOS technology was a 32-nm process, meaning that for this process the minimum channel length possible was 32 nm. Finally, we should note that the oxide thickness t_{ox} scales down with L_{min}. Thus, for a 0.13-μm technology, t_{ox} is 2.7 nm, but for the currently popular 65-nm technology, t_{ox} is about 2.2 nm.

[5]This name arises from the fact that $(\mu_n C_{ox})$ determines the transconductance of the MOSFET, as will be seen shortly.

The product of the process transconductance parameter k_n' and the transistor aspect ratio (W/L) is the **MOSFET transconductance parameter** k_n,

$$k_n = k_n'(W/L) \tag{5.12a}$$

or

$$k_n = (\mu_n C_{ox})(W/L) \tag{5.12b}$$

The MOSFET parameter k_n has the dimensions of A/V^2.

The third term in the expression of the channel conductance g_{DS} is the overdrive voltage v_{OV}. This is hardly surprising, since v_{OV} directly determines the magnitude of electron charge in the channel. As will be seen, v_{OV} is a very important circuit-design parameter. In this book, we will use v_{OV} and $v_{GS} - V_t$ interchangeably.

We conclude this subsection by noting that with v_{DS} kept small, the MOSFET behaves as a linear resistance r_{DS} whose value is controlled by the gate voltage v_{GS},

$$r_{DS} = \frac{1}{g_{DS}}$$

$$r_{DS} = \frac{1}{(\mu_n C_{ox})(W/L)v_{OV}} \tag{5.13a}$$

$$r_{DS} = \frac{1}{(\mu_n C_{ox})(W/L)(v_{GS} - V_t)} \tag{5.13b}$$

The operation of the MOSFET as a voltage-controlled resistance is further illustrated in Fig. 5.4, which is a sketch of i_D versus v_{DS} for various values of v_{GS}. Observe that the resistance is infinite for $v_{GS} \leq V_t$ and decreases as v_{GS} is increased above V_t. It is interesting to note that although v_{GS} is used as the parameter for the set of graphs in Fig. 5.4, the graphs in fact depend only on v_{OV} (and, of course, k_n).

The description above indicates that for the MOSFET to conduct, a channel has to be induced. Then, increasing v_{GS} above the threshold voltage V_t enhances the channel, hence the names **enhancement-mode operation** and **enhancement-type MOSFET**. Finally, we note that the current that leaves the source terminal (i_S) is equal to the current that enters the drain terminal (i_D), and the gate current $i_G = 0$.

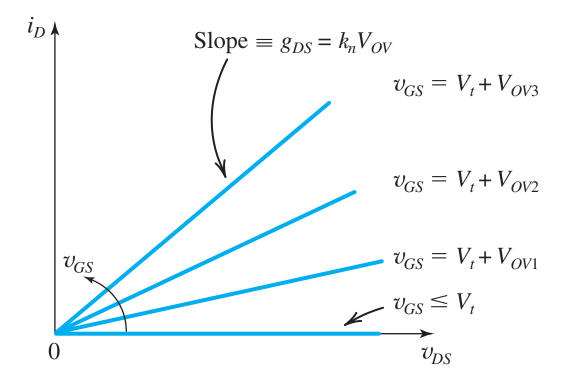


Figure 5.4 The i_D–v_{DS} characteristics of the MOSFET in Fig. 5.3 when the voltage applied between drain and source, v_{DS}, is kept small. The device operates as a linear resistance whose value is controlled by v_{GS}.

EXERCISE

5.1 A 0.18-μm fabrication process is specified to have $t_{ox} = 4$ nm, $\mu_n = 450\ \text{cm}^2/\text{V}\cdot\text{s}$, and $V_t = 0.5$ V. Find the value of the process transconductance parameter k_n'. For a MOSFET with minimum length fabricated in this process, find the required value of W so that the device exhibits a channel resistance r_{DS} of 1 kΩ at $v_{GS} = 1$ V.
Ans. 388 μA/V^2; 0.93 μm

5.1.5 Operation as v_{DS} Is Increased

We next consider the situation as v_{DS} is increased. For this purpose, let v_{GS} be held constant at a value greater than V_t; that is, let the MOSFET be operated at a constant overdrive voltage V_{OV}. Refer to Fig. 5.5, and note that v_{DS} appears as a voltage drop across the length of the channel. That is, as we travel along the channel from source to drain, the voltage (measured relative to the source) increases from zero to v_{DS}. Thus the voltage between the gate and points along the channel decreases from $v_{GS} = V_t + V_{OV}$ at the source end to $v_{GD} = v_{GS} - v_{DS} = V_t + V_{OV} - v_{DS}$ at the drain end. Since the channel depth depends on this voltage, and specifically on the amount by which this voltage exceeds V_t, we find that the channel is no longer of uniform depth; rather, the channel will take the tapered shape shown in Fig. 5.5, being deepest at the source end (where the depth is proportional to V_{OV}) and shallowest at the drain end[6] (where the depth is proportional to $V_{OV} - v_{DS}$). This point is further illustrated in Fig. 5.6.

As v_{DS} is increased, the channel becomes more tapered and its resistance increases correspondingly. Thus, the $i_D - v_{DS}$ curve does not continue as a straight line but bends as shown in Fig. 5.7. The equation describing this portion of the $i_D - v_{DS}$ curve can be easily derived by utilizing the information in Fig. 5.6. Specifically, note that the charge in the tapered channel is proportional to the channel cross-sectional area shown in Fig. 5.6(b). This area in turn can be easily seen as proportional to $\frac{1}{2}[V_{OV} + (V_{OV} - v_{DS})]$ or $\left(V_{OV} - \frac{1}{2}v_{DS}\right)$. Thus, the relationship between i_D and v_{DS} can be found by replacing V_{OV} in Eq. (5.7) by $\left(V_{OV} - \frac{1}{2}v_{DS}\right)$,

$$i_D = k_n'\left(\frac{W}{L}\right)\left(V_{OV} - \frac{1}{2}v_{DS}\right)v_{DS} \tag{5.14}$$

This relationship describes the semiparabolic portion of the $i_D - v_{DS}$ curve in Fig. 5.7. It applies to the entire segment down to $v_{DS} = 0$. Specifically, note that as v_{DS} is reduced, we can neglect $\frac{1}{2}v_{DS}$ relative to V_{OV} in the factor in parentheses, and the expression reduces to that in Eq. (5.7). The latter of course is an approximation and applies only for small v_{DS} (i.e., near the origin).

There is another useful interpretation of the expression in Eq. (5.14). From Fig. 5.6(a) we see that the average voltage along the channel is $\frac{1}{2}v_{DS}$. Thus, the average voltage that gives rise to channel charge and hence to i_D is no longer V_{OV} but $\left(V_{OV} - \frac{1}{2}v_{DS}\right)$, which is indeed the factor that appears in Eq. (5.14). Finally, we note that Eq. (5.14) is frequently written in the

[6]For simplicity, we do not show in Fig. 5.5 the depletion region. Physically speaking, it is the widening of the depletion region as a result of the increased v_{DS} that makes the channel shallower near the drain.

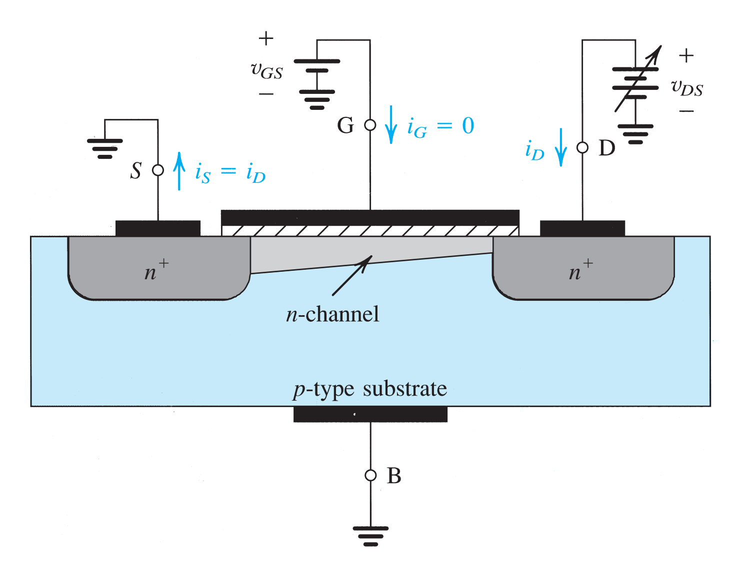


Figure 5.5 Operation of the enhancement NMOS transistor as v_{DS} is increased. The induced channel acquires a tapered shape, and its resistance increases as v_{DS} is increased. Here, v_{GS} is kept constant at a value $> V_t$; $v_{GS} = V_t + V_{OV}$.

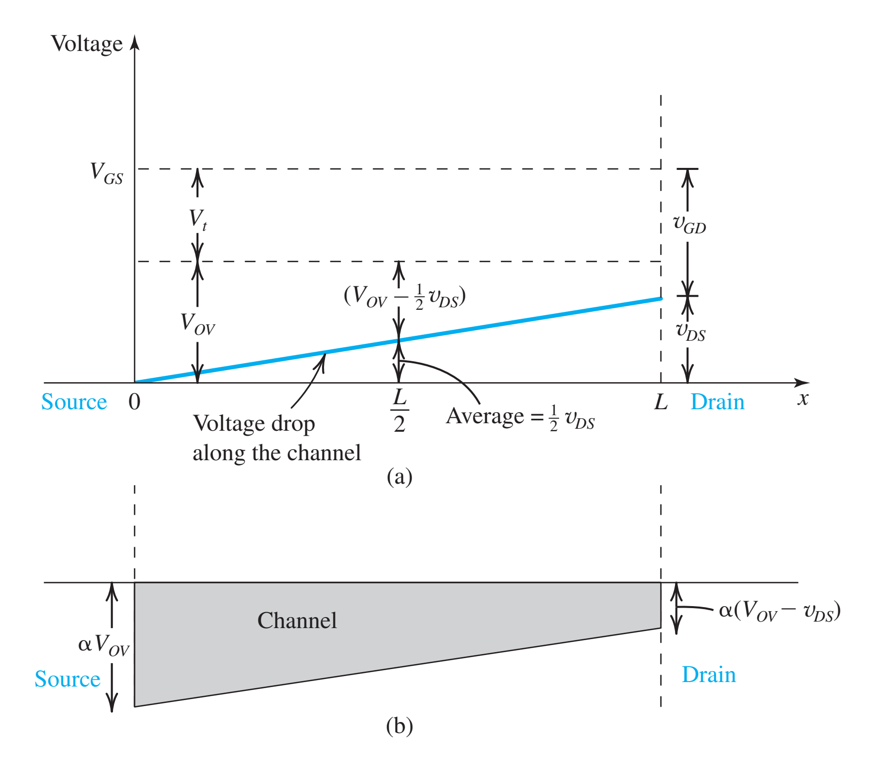


Figure 5.6 **(a)** For a MOSFET with $v_{GS} = V_t + V_{OV}$, application of v_{DS} causes the voltage drop along the channel to vary linearly, with an average value of $\frac{1}{2}v_{DS}$ at the midpoint. Since $v_{GD} > V_t$, the channel still exists at the drain end. **(b)** The channel shape corresponding to the situation in **(a)**. While the depth of the channel at the source end is still proportional to V_{OV}, that at the drain end is proportional to $(V_{OV} - v_{DS})$.

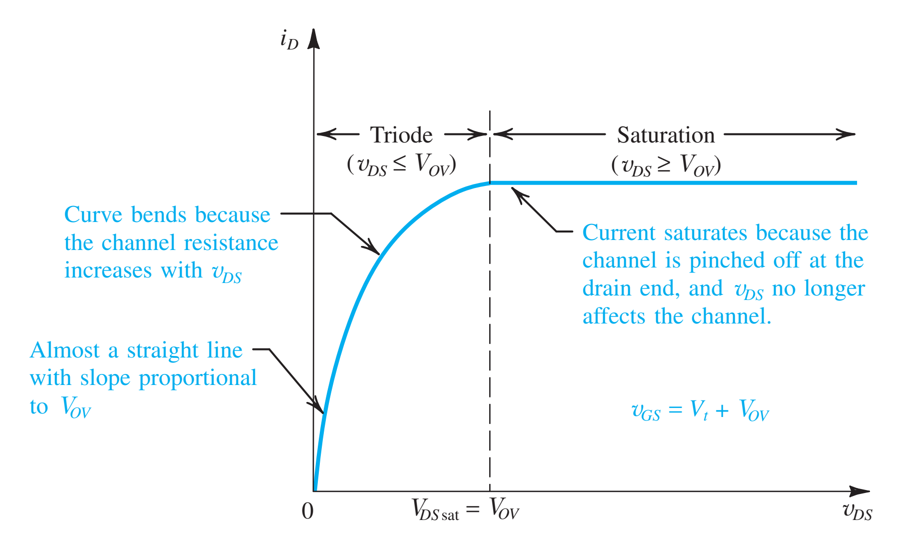


Figure 5.7 The drain current i_D versus the drain-to-source voltage v_{DS} for an enhancement-type NMOS transistor operated with $v_{GS} = V_t + V_{OV}$.

alternate form

$$i_D = k_n'\left(\frac{W}{L}\right)\left(V_{OV}v_{DS} - \frac{1}{2}v_{DS}^2\right) \tag{5.15}$$

Furthermore, for an arbitrary value of V_{OV}, we can replace V_{OV} by $(v_{GS} - V_t)$ and rewrite Eq. (5.15) as

$$i_D = k_n'\left(\frac{W}{L}\right)\left[(v_{GS} - V_t)v_{DS} - \frac{1}{2}v_{DS}^2\right] \tag{5.16}$$

5.1.6 Operation for $v_{DS} \geq V_{OV}$: Channel Pinch-Off and Current Saturation

The above description of operation assumed that even though the channel became tapered, it still had a finite (nonzero) depth at the drain end. This in turn is achieved by keeping v_{DS} sufficiently small that the voltage between the gate and the drain, v_{GD}, exceeds V_t. This is indeed the situation shown in Fig. 5.6(a). Note that for this situation to obtain, v_{DS} must not exceed V_{OV}, for as $v_{DS} = V_{OV}$, $v_{GD} = V_t$, and the channel depth at the drain end reduces to zero.

Figure 5.8 shows v_{DS} reaching V_{OV} and v_{GD} correspondingly reaching V_t. The zero depth of the channel at the drain end gives rise to the term **channel pinch-off**. Increasing v_{DS} beyond this value (i.e., $v_{DS} > V_{OV}$) has no effect on the channel shape and charge, and the current through the channel remains constant at the value reached for $v_{DS} = V_{OV}$. The drain current thus **saturates** at the value found by substituting $v_{DS} = V_{OV}$ in Eq. (5.14),

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)V_{OV}^2 \tag{5.17}$$

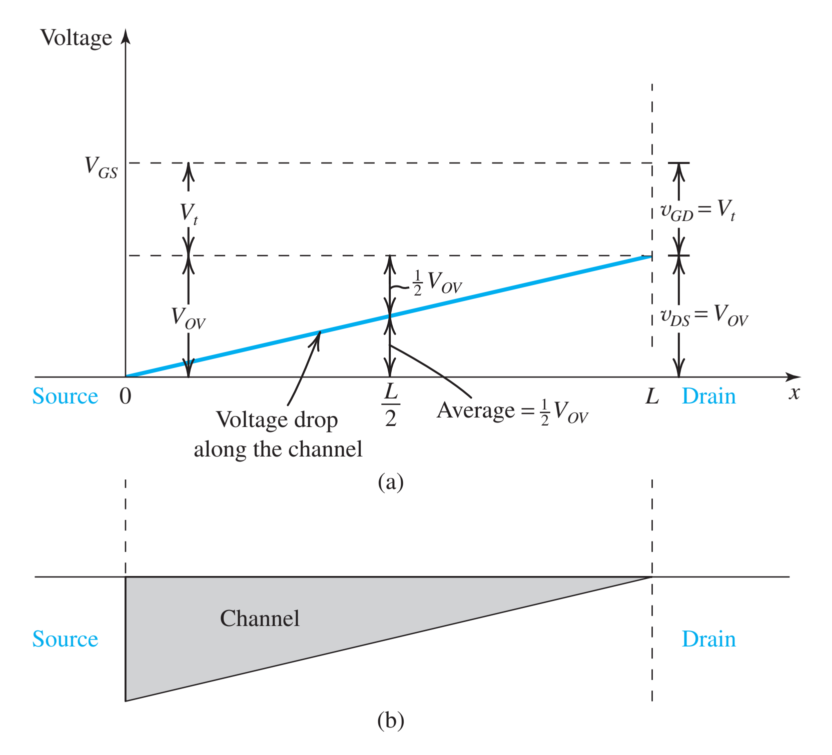


Figure 5.8 Operation of MOSFET with $v_{GS} = V_t + V_{OV}$, as v_{DS} is increased to V_{OV}. At the drain end, v_{GD} decreases to V_t and the channel depth at the drain end reduces to zero (pinch-off). At this point, the MOSFET enters the saturation mode of operation. Further increasing v_{DS} (beyond $V_{DS\text{sat}} = V_{OV}$) has no effect on the channel shape and i_D remains constant.

The MOSFET is then said to have entered the **saturation region** (or, equivalently, the saturation mode of operation). The voltage v_{DS} at which saturation occurs is denoted $V_{DS\text{sat}}$,

$$V_{DS\text{sat}} = V_{OV} = V_{GS} - V_t \tag{5.18}$$

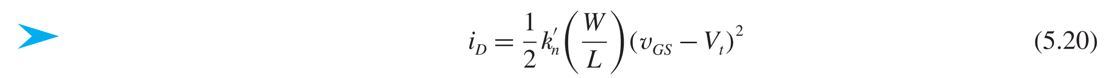
It should be noted that channel pinch-off does *not* mean channel blockage: Current continues to flow through the pinched-off channel, and the electrons that reach the drain end of the channel are accelerated through the depletion region that exists there (not shown in Fig. 5.5) and into the drain terminal. Any increase in v_{DS} above $V_{DS\text{sat}}$ appears as a voltage drop across the depletion region. Thus, both the current through the channel and the voltage drop across it remain constant in saturation.

The saturation portion of the $i_D - v_{DS}$ curve is, as expected, a horizontal straight line, as indicated in Fig. 5.7. Also indicated in Fig. 5.7 is the name of the region of operation obtained with a continuous (non-pinched-off) channel, the **triode region**. This name is a carryover from the days of vacuum-tube devices, whose operation a FET resembles.

Finally, we note that the expression for i_D in saturation can be generalized by replacing the constant overdrive voltage V_{OV} by a variable one, v_{OV}:

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)v_{OV}^2 \tag{5.19}$$

Also, v_{OV} can be replaced by $(v_{GS} - V_t)$ to obtain the alternate expression for saturation-mode i_D,



$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)(v_{GS} - V_t)^2 \tag{5.20}$$

Example 5.1

Consider a process technology for which $L_{\min} = 0.4\ \mu\text{m}$, $t_{ox} = 8$ nm, $\mu_n = 450\ \text{cm}^2/\text{V}\cdot\text{s}$, and $V_t = 0.7$ V.

(a) Find C_{ox} and k_n'.
(b) For a MOSFET with $W/L = 8\ \mu\text{m}/0.8\ \mu\text{m}$, calculate the values of V_{OV}, V_{GS}, and $V_{DS\min}$ needed to operate the transistor in the saturation region with a dc current $I_D = 100\ \mu\text{A}$.
(c) For the device in (b), find the values of V_{OV} and V_{GS} required to cause the device to operate as a 1000-Ω resistor for very small v_{DS}.

Solution

(a)

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = \frac{3.45 \times 10^{-11}}{8 \times 10^{-9}} = 4.32 \times 10^{-3}\ \text{F/m}^2$$
$$= 4.32\ \text{fF}/\mu\text{m}^2$$
$$k_n' = \mu_n C_{ox} = 450\ (\text{cm}^2/\text{V}\cdot\text{s}) \times 4.32\ (\text{fF}/\mu\text{m}^2)$$
$$= 450 \times 10^8\ (\mu\text{m}^2/\text{V}\cdot\text{s}) \times 4.32 \times 10^{-15}\ (\text{F}/\mu\text{m}^2)$$
$$= 194 \times 10^{-6}\ (\text{F/V}\cdot\text{s})$$
$$= 194\ \mu\text{A/V}^2$$

(b) For operation in the saturation region,

$$i_D = \frac{1}{2}k_n'\frac{W}{L}V_{OV}^2$$

Thus,

$$100 = \frac{1}{2} \times 194 \times \frac{8}{0.8}V_{OV}^2$$

which results in

$$V_{OV} = 0.32\ \text{V}$$

Thus,

$$V_{GS} = V_t + V_{OV} = 1.02\ \text{V}$$

and

$$V_{DS\min} = V_{OV} = 0.32\ \text{V}$$

(c) For the MOSFET in the triode region with v_{DS} very small,

$$r_{DS} = \frac{1}{k_n' \frac{W}{L} V_{OV}}$$

Thus

$$1000 = \frac{1}{194 \times 10^{-6} \times 10 \times V_{OV}}$$

which yields

$$V_{OV} = 0.52 \text{ V}$$

Thus,

$$V_{GS} = 1.22 \text{ V}$$

EXERCISES

5.2 For a 0.18-μm process technology for which $t_{ox} = 4$ nm and $\mu_n = 450$ cm^2/V·s, find C_{ox}, k_n', and the overdrive voltage V_{OV} required to operate a transistor having $W/L = 20$ in saturation with $I_D = 0.3$ mA. What is the minimum value of V_{DS} needed?
Ans. 8.6 fF/μm^2; 387 μA/V^2; 0.28 V; 0.28 V

D5.3 A circuit designer intending to operate a MOSFET in saturation is considering the effect of changing the device dimensions and operating voltages on the drain current I_D. Specifically, by what factor does I_D change in each of the following cases?
(a) The channel length is doubled.
(b) The channel width is doubled.
(c) The overdrive voltage is doubled.
(d) The drain-to-source voltage is doubled.
(e) Changes (a), (b), (c), and (d) are made simultaneously.
Which of these cases might cause the MOSFET to leave the saturation region?
Ans. 0.5; 2; 4; no change; 4; case (c) if v_{DS} is smaller than $2V_{OV}$

5.1.7 The *p*-Channel MOSFET

Figure 5.9(a) shows a cross-sectional view of a *p*-channel enhancement-type MOSFET. The structure is similar to that of the NMOS device except that here the substrate is *n* type and the source and the drain regions are p^+ type; that is, all semiconductor regions are reversed in polarity relative to their counterparts in the NMOS case. The PMOS and NMOS transistors are said to be *complementary* devices.

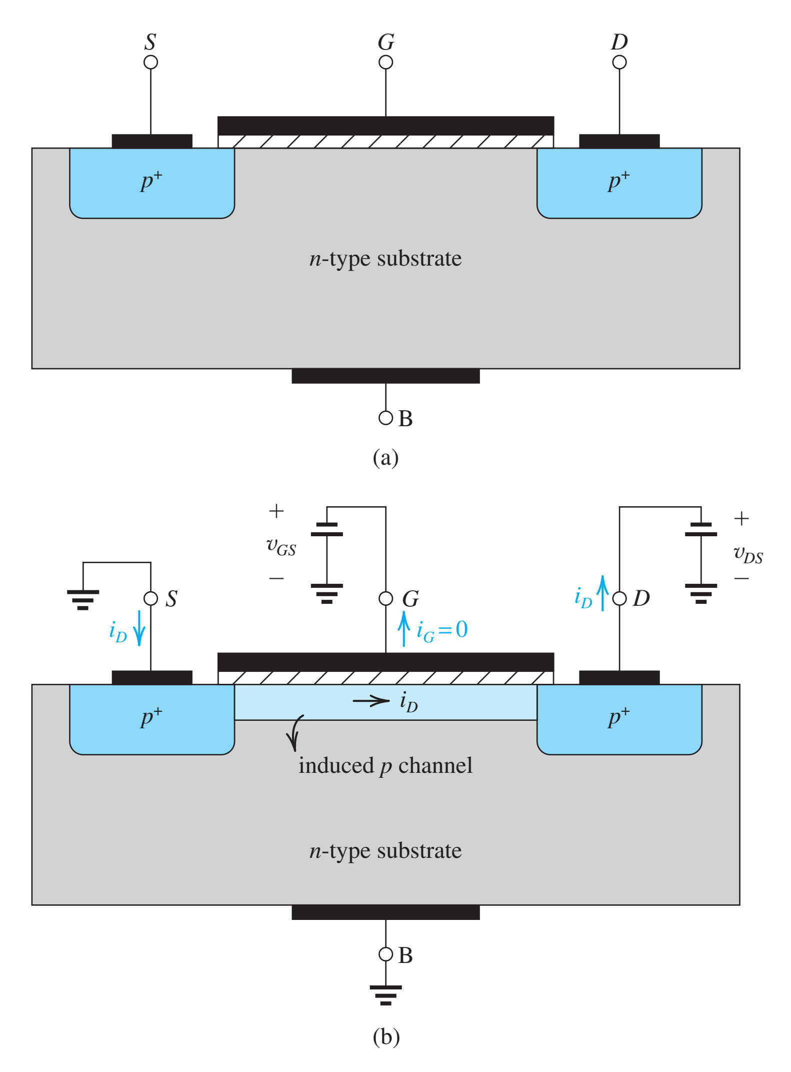


Figure 5.9 **(a)** Physical structure of the PMOS transistor. Note that it is similar to the NMOS transistor shown in Fig. 5.1(b) except that all semiconductor regions are reversed in polarity. **(b)** A negative voltage v_{GS} of magnitude greater than $|V_{tp}|$ induces a p channel, and a negative v_{DS} causes a current i_D to flow from source to drain.

To induce a channel for current flow between source and drain, a negative voltage is applied to the gate, that is, between gate and source, as indicated in Fig. 5.9(b). By increasing the magnitude of the negative v_{GS} beyond the magnitude of the threshold voltage V_{tp}, which by convention is negative, a p channel is established as shown in Fig. 5.9(b). This condition can be described as

$$v_{GS} \leq V_{tp}$$

or, to avoid dealing with negative signs,

$$|v_{GS}| \geq |V_{tp}|$$

Now, to cause a current i_D to flow in the p channel, a negative voltage v_{DS} is applied to the drain.[7] The current i_D is carried by holes and flows through the channel from source to drain. As we have done for the NMOS transistor, we define the process transconductance parameter for the PMOS device as

$$k_p' = \mu_p C_{ox}$$

where μ_p is the mobility of the holes in the induced p channel. Typically, $\mu_p = 0.25\,\mu_n$ to $0.5\,\mu_n$ and is process-technology dependent. The transistor transconductance parameter k_p is obtained by multiplying k_p' by the aspect ratio W/L,

$$k_p = k_p'(W/L)$$

The remainder of the description of the physical operation of the p-channel MOSFET follows that for the NMOS device, except of course for the sign reversals of all voltages. We will present the complete current–voltage characteristics of both NMOS and PMOS transistors in the next section.

PMOS technology originally dominated MOS integrated-circuit manufacturing, and the original microprocessors utilized PMOS transistors. As the technological difficulties of fabricating NMOS transistors were solved, NMOS completely supplanted PMOS. The main reason for this change is that electron mobility μ_n is higher by a factor of 2 to 4 than the hole mobility μ_p, resulting in NMOS transistors having greater gains and speeds of operation than PMOS devices. Subsequently, a technology was developed that permits the fabrication of both NMOS and PMOS transistors on the same chip. Appropriately called **complementary MOS**, or **CMOS**, this technology is currently the dominant electronics technology.

5.1.8 Complementary MOS or CMOS

As the name implies, complementary MOS technology employs MOS transistors of both polarities. Although CMOS circuits are somewhat more difficult to fabricate than NMOS, the availability of complementary devices makes possible many powerful circuit configurations. Indeed, at the present time CMOS is the most widely used of all the IC technologies. This statement applies to both analog and digital circuits. CMOS technology has virtually replaced designs based on NMOS transistors alone. Furthermore, by 2014 CMOS technology had taken over many applications that just a few years earlier were possible only with bipolar devices. Throughout this book, we will study many CMOS circuit techniques.

Figure 5.10 shows a cross section of a CMOS chip illustrating how the PMOS and NMOS transistors are fabricated. Observe that while the NMOS transistor is implemented directly in the p-type substrate, the PMOS transistor is fabricated in a specially created n region, known as an ***n* well**. The two devices are isolated from each other by a thick region of oxide that functions as an insulator. Not shown on the diagram are the connections made to the p-type body and to the n well. The latter connection serves as the body terminal for the PMOS transistor.

---

[7]If a positive voltage is applied to the drain, the *pn* junction between the drain region and the substrate will become forward biased, and the device will no longer operate as a MOSFET. Proper MOSFET operation is predicated on the *pn* junctions between the source and drain regions and the substrate being always reverse biased.

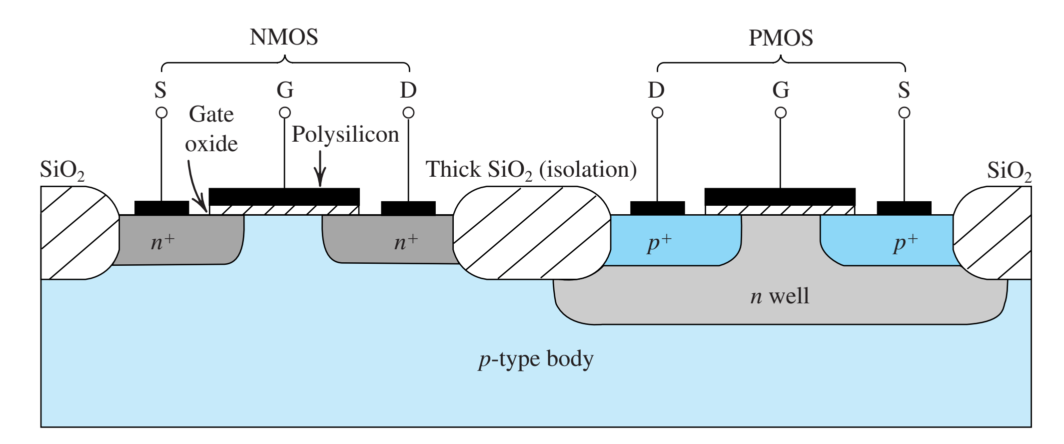


Figure 5.10 Cross section of a CMOS integrated circuit. Note that the PMOS transistor is formed in a separate *n*-type region, known as an *n* well. Another arrangement is also possible in which an *n*-type substrate (body) is used and the *n* device is formed in a *p* well. Not shown are the connections made to the *p*-type body and to the *n* well; the latter functions as the body terminal for the *p*-channel device.

5.1.9 Operating the MOS Transistor in the Subthreshold Region

The above description of the *n*-channel MOSFET operation implies that for $v_{GS} < V_t$, no current flows and the device is cut off. This is not entirely true, for it has been found that for values of v_{GS} smaller than but close to V_t, a small drain current flows. In this **subthreshold region** of operation, the drain current is exponentially related to v_{GS}, much like the i_C–v_{BE} relationship of a BJT, as will be shown in the next chapter.

Although in most applications the MOS transistor is operated with $v_{GS} > V_t$, there are special, but a growing number of, applications that make use of subthreshold operation. In Chapter 14, we will briefly consider subthreshold operation.

5.2 Current–Voltage Characteristics

Building on the physical foundation established in the previous section for the operation of the enhancement MOS transistor, in this section we present its complete current–voltage characteristics. These characteristics can be measured at dc or at low frequencies and thus are called static characteristics. The dynamic effects that limit the operation of the MOSFET at high frequencies and high switching speeds will be discussed in Chapter 10.

5.2.1 Circuit Symbol

Figure 5.11(a) shows the circuit symbol for the *n*-channel enhancement-type MOSFET. Observe that the spacing between the two vertical lines that represent the gate and the channel indicates the fact that the gate electrode is insulated from the body of the device. The polarity of the *p*-type substrate (body) and the *n* channel is indicated by the arrowhead on the line representing the body (B). This arrowhead also indicates the polarity of the transistor, namely, that it is an *n*-channel device.

Although the MOSFET is a symmetrical device, it is often useful in circuit design to designate one terminal as the source and the other as the drain (without having to write S and

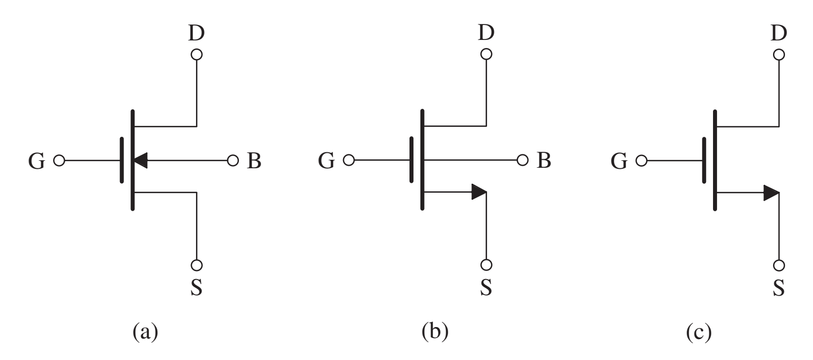


Figure 5.11 **(a)** Circuit symbol for the *n*-channel enhancement-type MOSFET. **(b)** Modified circuit symbol with an arrowhead on the source terminal to distinguish it from the drain and to indicate device polarity (i.e., *n* channel). **(c)** Simplified circuit symbol to be used when the source is connected to the body or when the effect of the body on device operation is unimportant.

D beside the terminals). This objective is achieved in the modified circuit symbol shown in Fig. 5.11(b). Here an arrowhead is placed on the source terminal, thus distinguishing it from the drain terminal. The arrowhead points in the normal direction of current flow and thus indicates the polarity of the device (i.e., *n* channel). Observe that in the modified symbol, there is no need to show the arrowhead on the body line. Although the circuit symbol of Fig. 5.11(b) clearly distinguishes the source from the drain, in practice it is the polarity of the voltage impressed across the device that determines source and drain; *the drain is always positive relative to the source in an n-channel FET.*

In applications where the source is connected to the body of the device, a further simplification of the circuit symbol is possible, as indicated in Fig. 5.11(c). This symbol is also used in applications when the effect of the body on circuit operation is not important, as will be seen later.

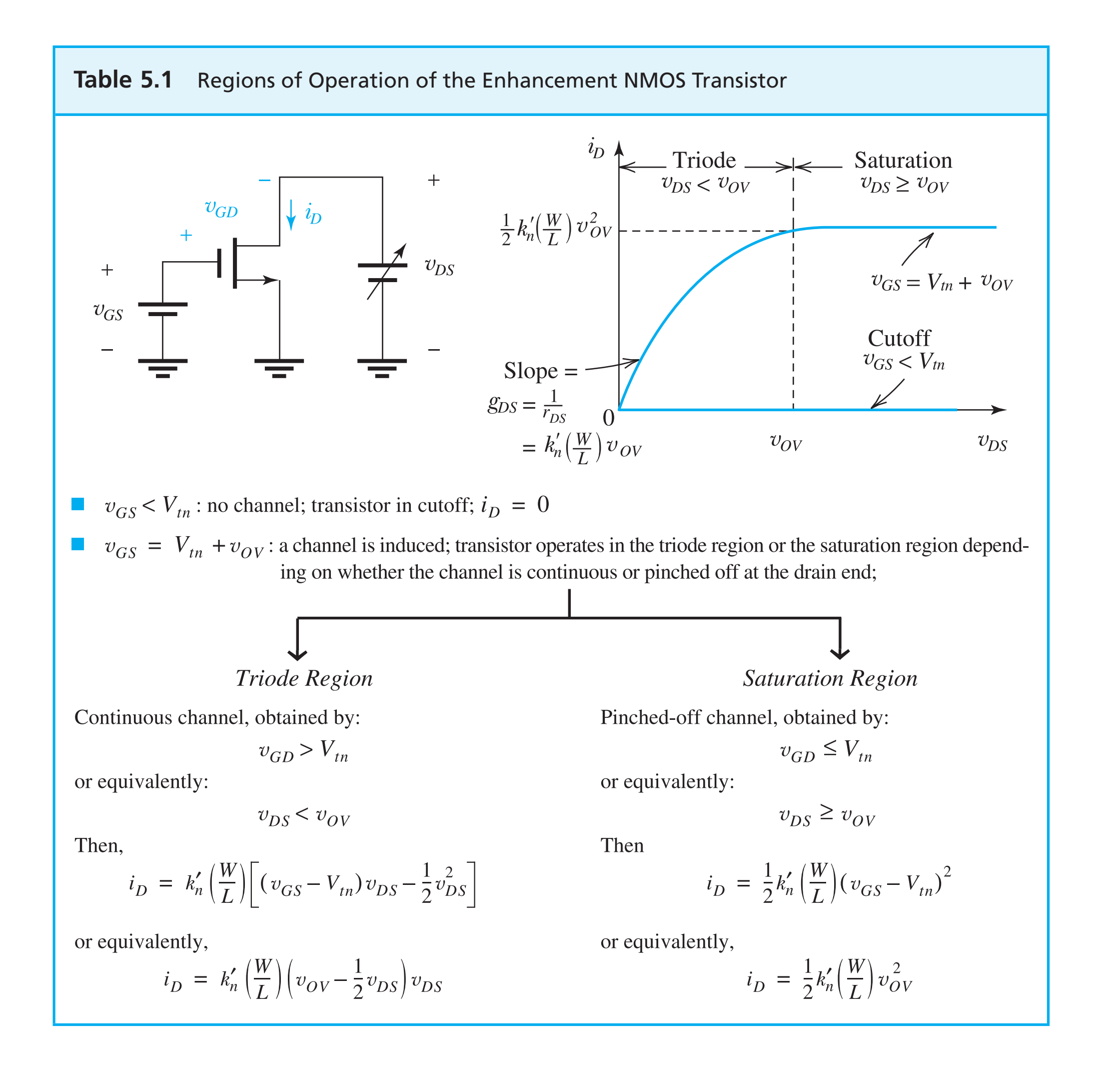
5.2.2 The i_D–v_{DS} Characteristics

Table 5.1 provides a compilation of the conditions and the formulas for the operation of the NMOS transistor in each of the three possible regions: the cutoff region, the triode region, and the saturation region. The first two are useful if the MOSFET is to be utilized as a switch. On the other hand, if the MOSFET is to be used to design an amplifier, it must be operated in the saturation region. The rationale for these choices will be addressed in Chapter 7.

At the top of Table 5.1 we show a circuit consisting of an NMOS transistor and two dc supplies providing v_{GS} and v_{DS}. This conceptual circuit can be used to measure the i_D–v_{DS} characteristic curves of the NMOS transistor. Each curve is measured by setting v_{GS} to a desired constant value, varying v_{DS}, and measuring the corresponding i_D. Two of these characteristic curves are shown in the accompanying diagram: one for $v_{GS} < V_{tn}$ and the other for $v_{GS} = V_{tn} + v_{OV}$. (Note that we now use V_{tn} to denote the threshold voltage of the NMOS transistor, to distinguish it from that of the PMOS transistor, denoted V_{tp}.)

As Table 5.1 shows, the boundary between the triode region and the saturation region is determined by whether v_{DS} is less or greater than the overdrive voltage v_{OV} at which the transistor is operating. An equivalent way to check for the region of operation is to examine the relative values of the drain and gate voltages. To operate in the triode region, the gate voltage must exceed the drain voltage by at least V_{tn} volts, which ensures that the channel remains continuous (not pinched off). On the other hand, to operate in saturation, the channel must be

Table 5.1 Regions of Operation of the Enhancement NMOS Transistor



- $v_{GS} < V_{tn}$: no channel; transistor in cutoff; $i_D = 0$
- $v_{GS} = V_{tn} + v_{OV}$: a channel is induced; transistor operates in the triode region or the saturation region depending on whether the channel is continuous or pinched off at the drain end;

Triode Region

Continuous channel, obtained by:

$$v_{GD} > V_{tn}$$

or equivalently:

$$v_{DS} < v_{OV}$$

Then,

$$i_D = k_n'\left(\frac{W}{L}\right)\left[(v_{GS} - V_{tn})v_{DS} - \frac{1}{2}v_{DS}^2\right]$$

or equivalently,

$$i_D = k_n'\left(\frac{W}{L}\right)\left(v_{OV} - \frac{1}{2}v_{DS}\right)v_{DS}$$

Saturation Region

Pinched-off channel, obtained by:

$$v_{GD} \leq V_{tn}$$

or equivalently:

$$v_{DS} \geq v_{OV}$$

Then

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)(v_{GS} - V_{tn})^2$$

or equivalently,

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)v_{OV}^2$$

pinched off at the drain end; pinch-off is achieved here by keeping v_D higher than $v_G - V_{tn}$, that is, not allowing v_D to fall below v_G by more than V_{tn} volts. The graphical construction of Fig. 5.12 should serve to remind the reader of these conditions.

A set of $i_D - v_{DS}$ characteristics for the NMOS transistor is shown in Fig. 5.13. Observe that each graph is obtained by setting v_{GS} above V_{tn} by a specific value of overdrive voltage, denoted V_{OV1}, V_{OV2}, V_{OV3}, and V_{OV4}. This in turn is the value of v_{DS} at which the corresponding graph saturates, and the value of the resulting saturation current is directly determined by the value of v_{OV}, namely, $\frac{1}{2}k_n V_{OV1}^2$, $\frac{1}{2}k_n V_{OV2}^2, \ldots$ The reader is advised to commit to memory both the structure of these graphs and the coordinates of the saturation points.

Finally, observe that the boundary between the triode and the saturation regions, that is, the locus of the saturation points, is a parabolic curve described by

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)v_{DS}^2$$

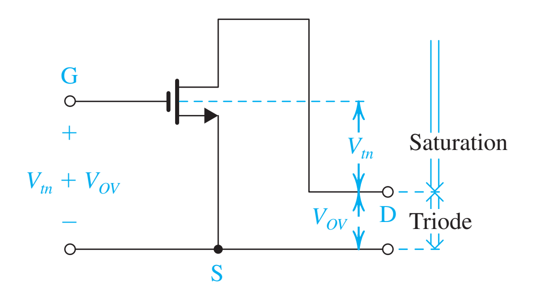


Figure 5.12 The relative levels of the terminal voltages of the enhancement NMOS transistor for operation in the triode region and in the saturation region.

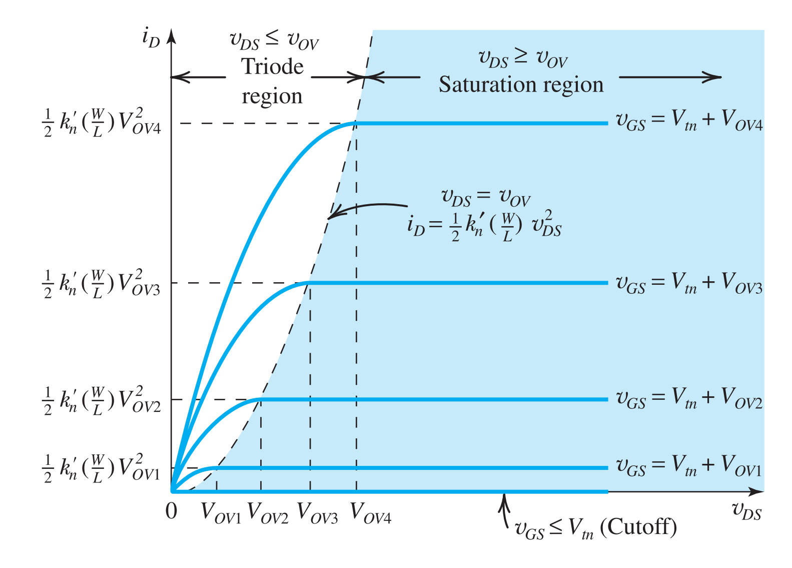


Figure 5.13 The $i_D - v_{DS}$ characteristics for an enhancement-type NMOS transistor.

5.2.3 The i_D–v_{GS} Characteristic

When the MOSFET is used to design an amplifier, it is operated in the saturation region. As Fig. 5.13 indicates, in saturation the drain current is constant determined by v_{GS} (or v_{OV}) and is independent of v_{DS}. That is, the MOSFET operates as a constant-current source where the value of the current is determined by v_{GS}. In effect, then, the MOSFET operates as a voltage-controlled current source with the control relationship described by

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)(v_{GS} - V_{tn})^2 \tag{5.21}$$

or in terms of v_{OV},

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)v_{OV}^2 \tag{5.22}$$

This is the relationship that underlies the application of the MOSFET as an amplifier. That it is nonlinear should be of concern to those interested in designing linear amplifiers. Nevertheless, in Chapter 7, we will see how one can obtain linear amplification from this nonlinear control or transfer characteristic.

Figure 5.14 shows the i_D–v_{GS} characteristic of an NMOS transistor operating in saturation. Note that if we are interested in a plot of i_D versus v_{OV}, we simply shift the origin to the point $v_{GS} = V_{tn}$.

The view of the MOSFET in the saturation region as a voltage-controlled current source is illustrated by the equivalent-circuit representation shown in Fig. 5.15. For reasons that will become apparent shortly, the circuit in Fig. 5.15 is known as a **large-signal equivalent circuit**. Note that the current source is ideal, with an infinite output resistance representing the independence, in saturation, of i_D from v_{DS}. This, of course, has been assumed in the idealized model of device operation utilized thus far. We are about to rectify an important shortcoming of this model. First, however, we present an example.

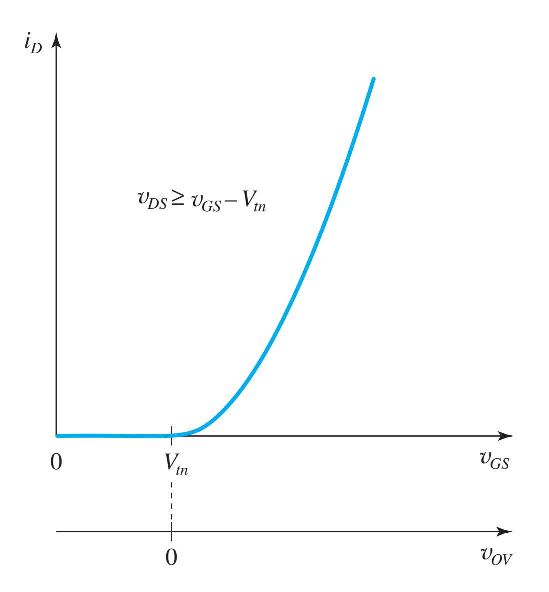


Figure 5.14 The i_D–v_{GS} characteristic of an NMOS transistor operating in the saturation region. The i_D–v_{OV} characteristic can be obtained by simply relabeling the horizontal axis, that is, shifting the origin to the point $v_{GS} = V_{tn}$.

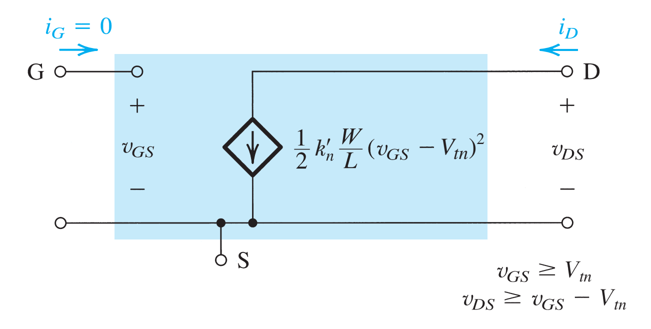


Figure 5.15 Large-signal, equivalent-circuit model of an n-channel MOSFET operating in the saturation region.

Example 5.2

Consider an NMOS transistor fabricated in a 0.18-μm process with $L = 0.18$ μm and $W = 2$ μm. The process technology is specified to have $C_{ox} = 8.6$ fF/μm^2, $\mu_n = 450$ cm^2/V·s, and $V_{tn} = 0.5$ V.

(a) Find V_{GS} and V_{DS} that result in the MOSFET operating at the edge of saturation with $I_D = 100$ μA.
(b) If V_{GS} is kept constant, find V_{DS} that results in $I_D = 50$ μA.
(c) To investigate the use of the MOSFET as a linear amplifier, let it be operating in saturation with $V_{DS} = 0.3$ V. Find the change in i_D resulting from v_{GS} changing from 0.7 V by +0.01 V and by −0.01 V.

Solution

First we determine the process transconductance parameter k_n',

$$\begin{aligned} k_n' &= \mu_n C_{ox} \\ &= 450 \times 10^{-4} \times 8.6 \times 10^{-15} \times 10^{12} \text{ A/V}^2 \\ &= 387 \text{ μA/V}^2 \end{aligned}$$

and the transistor transconductance parameter k_n,

$$\begin{aligned} k_n &= k_n'\left(\frac{W}{L}\right) \\ &= 387\left(\frac{2}{0.18}\right) = 4.3 \text{ mA/V}^2 \end{aligned}$$

(a) With the transistor operating in saturation,

$$I_D = \frac{1}{2}k_n V_{OV}^2$$

Thus,

$$100 = \frac{1}{2} \times 4.3 \times 10^3 \times V_{OV}^2$$

which results in

$$V_{OV} = 0.22 \text{ V}$$

Thus,

$$V_{GS} = V_{tn} + V_{OV} = 0.5 + 0.22 = 0.72 \text{ V}$$

and since operation is at the edge of saturation,

$$V_{DS} = V_{OV} = 0.22 \text{ V}$$

Example 5.2 *continued*

(b) With V_{GS} kept constant at 0.72 V and I_D reduced from the value obtained at the edge of saturation, the MOSFET will now be operating in the triode region, thus

$$I_D = k_n\left[V_{OV}V_{DS} - \frac{1}{2}V_{DS}^2\right]$$

$$50 = 4.3 \times 10^3\left[0.22V_{DS} - \frac{1}{2}V_{DS}^2\right]$$

which can be rearranged to the form

$$V_{DS}^2 - 0.44V_{DS} + 0.023 = 0$$

This quadratic equation has two solutions

$$V_{DS} = 0.06 \text{ V} \quad \text{and} \quad V_{DS} = 0.39 \text{ V}$$

The second answer is greater than V_{OV} and thus is physically meaningless, since we know that the transistor is operating in the triode region. Thus we have

$$V_{DS} = 0.06 \text{ V}$$

(c) For $v_{GS} = 0.7$ V, $V_{OV} = 0.2$ V, and since $V_{DS} = 0.3$ V, the transistor is operating in saturation and

$$\begin{aligned} I_D &= \frac{1}{2}k_n V_{OV}^2 \\ &= \frac{1}{2} \times 4300 \times 0.04 \\ &= 86 \text{ µA} \end{aligned}$$

Now for $v_{GS} = 0.710$ V, $v_{OV} = 0.21$ V and

$$i_D = \frac{1}{2} \times 4300 \times 0.21^2 = 94.8 \text{ µA}$$

and for $v_{GS} = 0.690$ V, $v_{OV} = 0.19$ V, and

$$i_D = \frac{1}{2} \times 4300 \times 0.19^2 = 77.6 \text{ µA}$$

Thus, with $\Delta v_{GS} = +0.01$ V, $\Delta i_D = 8.8$ µA; and for $\Delta v_{GS} = -0.01$ V, $\Delta i_D = -8.4$ µA.

We conclude that the two changes are almost equal, an indication of almost-linear operation when the changes in v_{GS} are kept small. This is just a preview of the "small-signal operation" of the MOSFET studied in Chapter 7.

EXERCISES

5.4 An NMOS transistor is operating at the edge of saturation with an overdrive voltage V_{OV} and a drain current I_D. If V_{OV} is doubled, and we must maintain operation at the edge of saturation, what should V_{DS} be changed to? What value of drain current results?
Ans. $2V_{OV}$; $4I_D$

5.5 An n-channel MOSFET operating with $V_{OV} = 0.5$ V exhibits a linear resistance $r_{DS} = 1$ kΩ when v_{DS} is very small. What is the value of the device transconductance parameter k_n? What is the value of the current I_D obtained when v_{DS} is increased to 0.5 V? and to 1 V?
Ans. 2 mA/V^2; 0.25 mA; 0.25 mA

5.2.4 Finite Output Resistance in Saturation

Equation (5.21) and the corresponding large-signal equivalent circuit in Fig. 5.15, as well as the graphs in Fig. 5.13, indicate that in saturation, i_D is independent of v_{DS}. Thus, a change Δv_{DS} in the drain-to-source voltage causes a zero change in i_D, which implies that the incremental resistance looking into the drain of a saturated MOSFET is infinite. This, however, is an idealization based on the premise that once the channel is pinched off at the drain end, further increases in v_{DS} have no effect on the channel's shape. But, in practice, increasing v_{DS} beyond v_{OV} does affect the channel somewhat. Specifically, as v_{DS} is increased, the channel pinch-off point is moved slightly away from the drain, toward the source. This is illustrated in Fig. 5.16, from which we note that the voltage across the channel remains constant at v_{OV}, and the additional voltage applied to the drain appears as a voltage drop across the narrow depletion region between the end of the channel and the drain region. This voltage accelerates the electrons that reach the drain end of the channel and sweeps them across the depletion region into the drain. Note, however, that (with depletion-layer widening) the channel length is in effect reduced, from L to $L - \Delta L$, a phenomenon known as **channel-length modulation**. Now, since i_D is inversely proportional to the channel length (Eq. 5.21), i_D increases with v_{DS}.

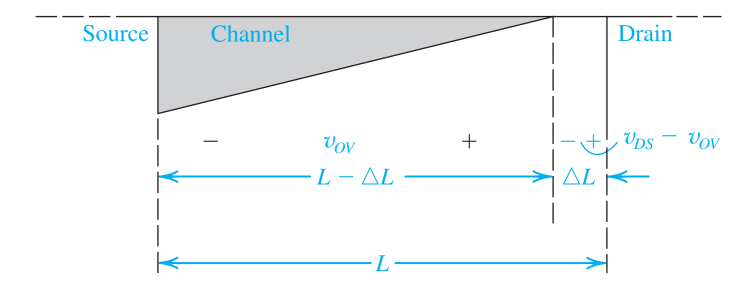


Figure 5.16 Increasing v_{DS} beyond v_{DSsat} causes the channel pinch-off point to move slightly away from the drain, thus reducing the effective channel length (by ΔL).

This effect can be accounted for in the expression for i_D by including a factor $1+\lambda(v_{DS}-v_{OV})$ or, for simplicity, $(1+\lambda v_{DS})$,

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)(v_{GS}-V_{tn})^2(1+\lambda v_{DS}) \tag{5.23}$$

Here λ is a device parameter having the units of reciprocal volts $\left(V^{-1}\right)$. The value of λ depends both on the process technology used to fabricate the device and on the channel length L that the circuit designer selects. Specifically, the value of λ is much larger for newer submicron technologies than for older technologies. This makes intuitive sense: Newer technologies have very short channels, and are thus much more greatly impacted by the channel-length modulation effect. Also, for a given process technology, λ is inversely proportional to L.

A typical set of i_D–v_{DS} characteristics showing the effect of channel-length modulation is displayed in Fig. 5.17. The observed linear dependence of i_D on v_{DS} in the saturation region is represented in Eq. (5.23) by the factor $(1+\lambda v_{DS})$. From Fig. 5.17 we observe that when the straight-line i_D–v_{DS} characteristics are extrapolated, they intercept the v_{DS} axis at the point, $v_{DS}=-V_A$, where V_A is a positive voltage. Equation (5.23), however, indicates that $i_D=0$ at $v_{DS}=-1/\lambda$. It follows that

$$V_A = \frac{1}{\lambda}$$

and thus V_A is a device parameter with the dimensions of V. For a given process, V_A is proportional to the channel length L that the designer selects for a MOSFET. We can isolate the dependence of V_A on L by expressing it as

$$V_A = V_A'L$$

where V_A' is entirely process-technology dependent, with the dimensions of volts per micron. Typically, V_A' falls in the range of 5 V/μm to 50 V/μm. The voltage V_A is usually referred to as the Early voltage, after J. M. Early, who discovered a similar phenomenon for the BJT (Chapter 6).

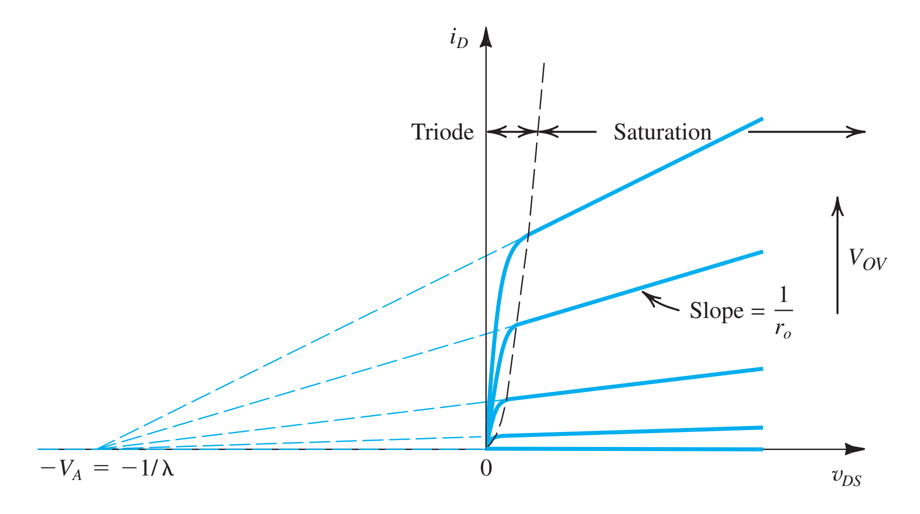


Figure 5.17 Effect of v_{DS} on i_D in the saturation region. The MOSFET parameter V_A depends on the process technology and, for a given process, is proportional to the channel length L.

Equation (5.23) indicates that when channel-length modulation is taken into account, the saturation values of i_D depend on v_{DS}. Thus, for a given v_{GS}, a change Δv_{DS} yields a corresponding change Δi_D in the drain current i_D. It follows that the output resistance of the current source representing i_D in saturation is no longer infinite. Defining the output resistance r_o as[8]

$$r_o \equiv \left[\frac{\partial i_D}{\partial v_{DS}}\right]^{-1}_{v_{GS}\ \text{constant}} \tag{5.24}$$

and using Eq. (5.23) results in

$$r_o = \left[\lambda \frac{k_n'}{2}\frac{W}{L}(V_{GS}-V_{tn})^2\right]^{-1} \tag{5.25}$$

which can be written as

$$r_o = \frac{1}{\lambda I_D} \tag{5.26}$$

or, equivalently,

$$r_o = \frac{V_A}{I_D'} \tag{5.27}$$

where I_D' is the drain current *without* channel-length modulation taken into account; that is,

$$I_D' = \frac{1}{2}k_n'\frac{W}{L}(V_{GS}-V_{tn})^2 \tag{5.27'}$$

Thus the output resistance is inversely proportional to the drain current.[9] Finally, we show in Fig. 5.18 the large-signal, equivalent-circuit model incorporating r_o.

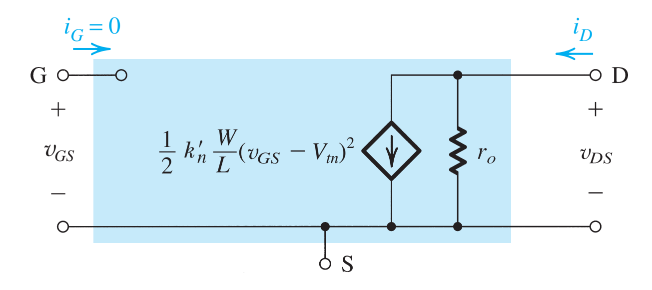


Figure 5.18 Large-signal, equivalent-circuit model of the *n*-channel MOSFET in saturation, incorporating the output resistance r_o. The output resistance models the linear dependence of i_D on v_{DS} and is given by Eq. (5.27).

[8] In this book we use r_o to denote the output resistance in saturation, and r_{DS} to denote the drain-to-source resistance in the triode region, for small v_{DS}.

[9] In applying Eq. (5.27) we will usually drop the prime on I_D and simply use $r_o = V_A/I_D$ where I_D is the drain current without channel-length modulation.

EXERCISE

5.6 An NMOS transistor is fabricated in a 0.4-μm process having $\mu_n C_{ox} = 200\ \mu\text{A/V}^2$ and $V_A' = 50$ V/μm of channel length. If $L = 0.8$ μm and $W = 16$ μm, find V_A and λ. Find the value of I_D that results when the device is operated with an overdrive voltage $V_{OV} = 0.5$ V and $V_{DS} = 1$ V. Also, find the value of r_o at this operating point. If V_{DS} is increased by 2 V, what is the corresponding change in I_D?
Ans. 40 V; 0.025 V^{-1}; 0.51 mA; 80 kΩ; 0.025 mA

5.2.5 Characteristics of the *p*-Channel MOSFET

The circuit symbol for the *p*-channel enhancement-type MOSFET is shown in Fig. 5.19(a). Figure 5.19(b) shows a modified circuit symbol in which an arrowhead pointing in the normal direction of current flow is included on the source terminal. For the case where the source is connected to the substrate, the simplified symbol of Fig. 5.19(c) is usually used.

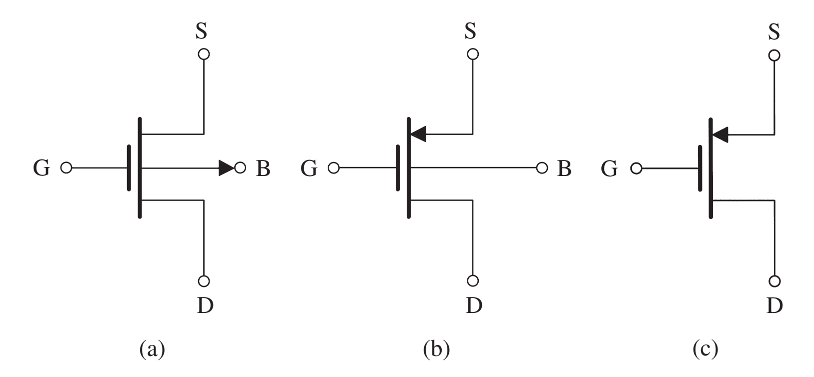


Figure 5.19 **(a)** Circuit symbol for the *p*-channel enhancement-type MOSFET. **(b)** Modified symbol with an arrowhead on the source lead. **(c)** Simplified circuit symbol for the case where the source is connected to the body.

The regions of operation of the PMOS transistor and the corresponding conditions and expression for i_D are shown in Table 5.2. Observe that the equations are written in a way that emphasizes physical intuition and avoids the confusion of negative signs. Thus while V_{tp} is by convention negative, we use $|V_{tp}|$, and the voltages v_{SG} and v_{SD} are positive. Also, in all of our circuit diagrams we will always draw *p*-channel devices with their sources on top so that current flows from top to bottom. Finally, we note that PMOS devices also suffer from the channel-length modulation effect. This can be taken into account by including a factor $(1 + |\lambda| v_{SD})$ in the saturation-region expression for i_D as follows

$$i_D = \frac{1}{2} k_p' \left(\frac{W}{L}\right) \left(v_{SG} - |V_{tp}|\right)^2 (1 + |\lambda| v_{SD}) \tag{5.28}$$

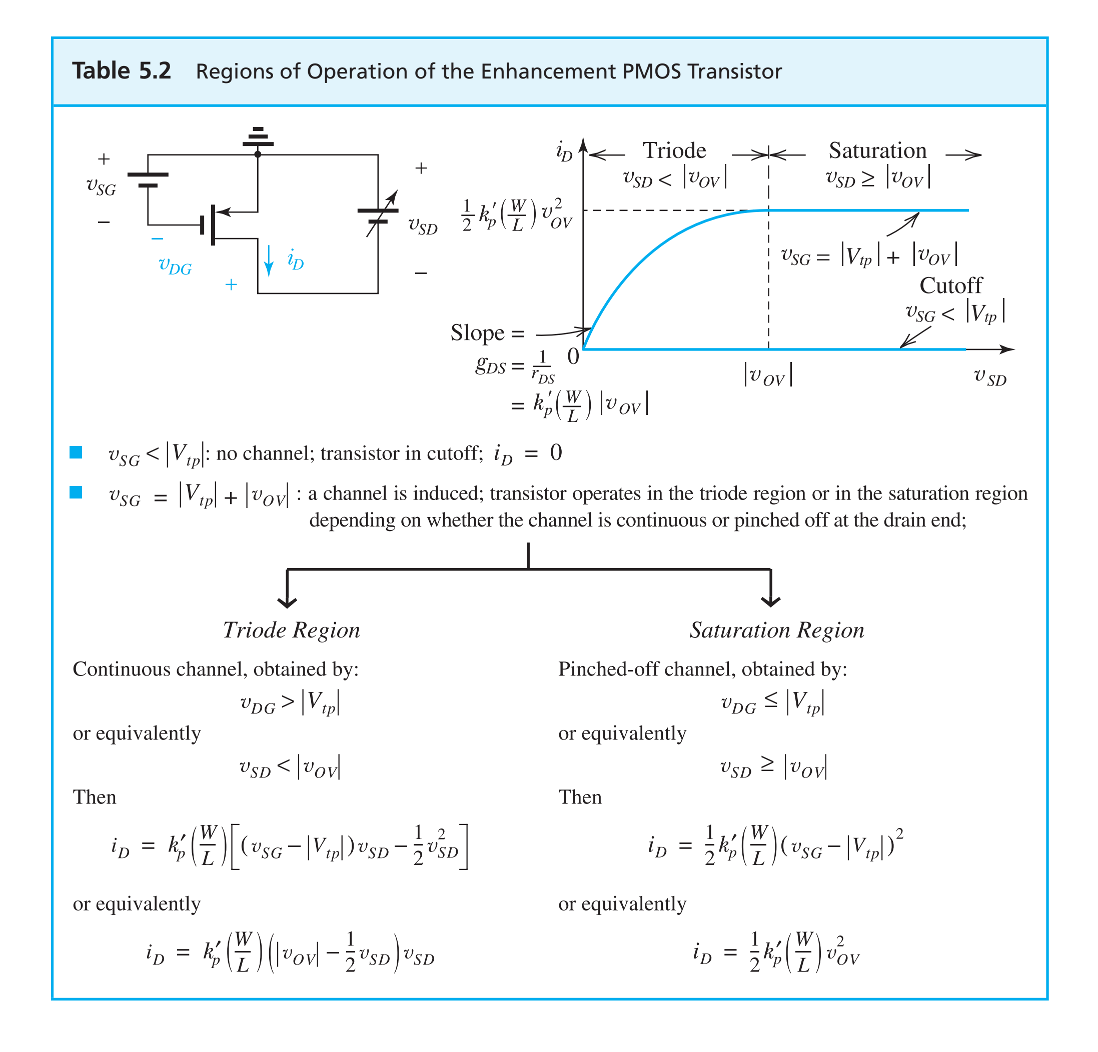


Table 5.2 Regions of Operation of the Enhancement PMOS Transistor

- $v_{SG} < |V_{tp}|$: no channel; transistor in cutoff; $i_D = 0$
- $v_{SG} = |V_{tp}| + |v_{OV}|$: a channel is induced; transistor operates in the triode region or in the saturation region depending on whether the channel is continuous or pinched off at the drain end;

Triode Region

Continuous channel, obtained by:

$$v_{DG} > |V_{tp}|$$

or equivalently

$$v_{SD} < |v_{OV}|$$

Then

$$i_D = k_p'\left(\frac{W}{L}\right)\left[(v_{SG} - |V_{tp}|)v_{SD} - \frac{1}{2}v_{SD}^2\right]$$

or equivalently

$$i_D = k_p'\left(\frac{W}{L}\right)\left(|v_{OV}| - \frac{1}{2}v_{SD}\right)v_{SD}$$

Saturation Region

Pinched-off channel, obtained by:

$$v_{DG} \leq |V_{tp}|$$

or equivalently

$$v_{SD} \geq |v_{OV}|$$

Then

$$i_D = \frac{1}{2}k_p'\left(\frac{W}{L}\right)(v_{SG} - |V_{tp}|)^2$$

or equivalently

$$i_D = \frac{1}{2}k_p'\left(\frac{W}{L}\right)v_{OV}^2$$

or equivalently

$$i_D = \frac{1}{2}k_p'\left(\frac{W}{L}\right)(v_{SG} - |V_{tp}|)^2\left(1 + \frac{v_{SD}}{|V_A|}\right) \tag{5.29}$$

where λ and V_A (the Early voltage for the PMOS transistor) are by convention negative quantities, hence we use $|\lambda|$ and $|V_A|$.

Finally, we should note that for a given CMOS fabrication process λ_n and $|\lambda_p|$ are generally not equal, and similarly for V_{An} and $|V_{Ap}|$.

To recap, to turn a PMOS transistor on, the gate voltage has to be made lower than that of the source by at least $|V_{tp}|$. To operate in the triode region, the drain voltage has to exceed that of the gate by at least $|V_{tp}|$; otherwise, the PMOS operates in saturation. Finally, Fig. 5.20 provides a pictorial representation of these operating conditions.

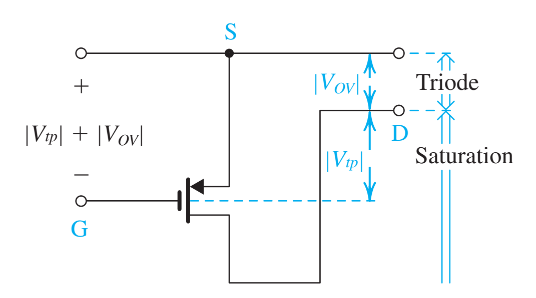


Figure 5.20 The relative levels of the terminal voltages of the enhancement-type PMOS transistor for operation in the triode region and in the saturation region.

EXERCISE

5.7 The PMOS transistor shown in Fig. E5.7 has $V_{tp} = -1$ V, $k'_p = 60\ \mu A/V^2$, and $W/L = 10$.

(a) Find the range of V_G for which the transistor conducts.
(b) In terms of V_G, find the range of V_D for which the transistor operates in the triode region.
(c) In terms of V_G, find the range of V_D for which the transistor operates in saturation.
(d) Neglecting channel-length modulation (i.e., assuming $\lambda = 0$), find the values of $|V_{OV}|$ and V_G and the corresponding range of V_D to operate the transistor in the saturation mode with $I_D = 75\ \mu A$.
(e) If $\lambda = -0.02\ V^{-1}$, find the value of r_o corresponding to the overdrive voltage determined in (d).
(f) For $\lambda = -0.02\ V^{-1}$ and for the value of V_{OV} determined in (d), find I_D at $V_D = +3$ V and at $V_D = 0$ V; hence, calculate the value of the apparent output resistance in saturation. Compare to the value found in (e).

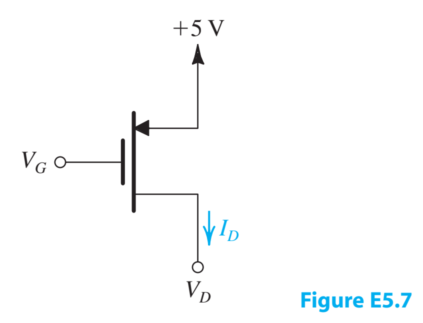


Figure E5.7

Ans. (a) $V_G \le +4$ V; (b) $V_D \ge V_G + 1$; (c) $V_D \le V_G + 1$; (d) 0.5 V, 3.5 V, ≤4.5 V; (e) 0.67 MΩ; (f) 78 μA, 82.5 μA, 0.67 MΩ (same)

5.3 MOSFET Circuits at DC

Having studied the current–voltage characteristics of MOSFETs, we now consider circuits in which only dc voltages and currents are of concern. Specifically, we shall present a series of design and analysis examples of MOSFET circuits at dc. The objective is to instill in the

reader a familiarity with the device and the ability to perform MOSFET circuit analysis both rapidly and effectively.

In the following examples, to keep matters simple and thus focus attention on the essence of MOSFET circuit operation, we will generally neglect channel-length modulation; that is, we will assume $\lambda = 0$. We will find it convenient to work in terms of the overdrive voltage; $V_{OV} = V_{GS} - V_{tn}$ for NMOS and $|V_{OV}| = V_{SG} - |V_{tp}|$ for PMOS.

Example 5.3

Design the circuit of Fig. 5.21: that is, determine the values of R_D and R_S so that the transistor operates at $I_D = 0.4$ mA and $V_D = +0.5$ V. The NMOS transistor has $V_t = 0.7$ V, $\mu_n C_{ox} = 100\ \mu\text{A/V}^2$, $L = 1\ \mu$m, and $W = 32\ \mu$m. Neglect the channel-length modulation effect (i.e., assume that $\lambda = 0$).

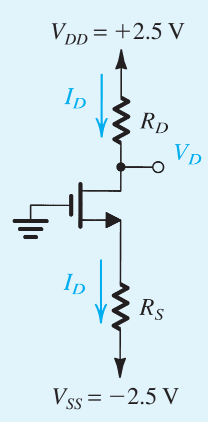


Figure 5.21 Circuit for Example 5.3.

Solution

To establish a dc voltage of +0.5 V at the drain, we must select R_D as follows:

$$R_D = \frac{V_{DD} - V_D}{I_D}$$
$$= \frac{2.5 - 0.5}{0.4} = 5\ \text{k}\Omega$$

To determine the value required for R_S, we need to know the voltage at the source, which can be easily found if we know V_{GS}. This in turn can be determined from V_{OV}. Toward that end, we note that since $V_D = 0.5$ V is greater than V_G, the NMOS transistor is operating in the saturation region, and we can use the saturation-region expression of i_D to determine the required value of V_{OV},

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}V_{OV}^2$$

Then substituting $I_D = 0.4\text{ mA} = 400\ \mu\text{A}$, $\mu_n C_{ox} = 100\ \mu\text{A/V}^2$, and $W/L = 32/1$ gives

$$400 = \frac{1}{2} \times 100 \times \frac{32}{1} V_{OV}^2$$

Example 5.3 *continued*

which results in

$$V_{OV} = 0.5 \text{ V}$$

Thus,

$$V_{GS} = V_t + V_{OV} = 0.7 + 0.5 = 1.2 \text{ V}$$

Referring to Fig. 5.21, we note that the gate is at ground potential. Thus, the source must be at –1.2 V, and the required value of R_S can be determined from

$$R_S = \frac{V_S - V_{SS}}{I_D}$$

$$= \frac{-1.2 - (-2.5)}{0.4} = 3.25 \text{ k}\Omega$$

EXERCISE

D5.8 Redesign the circuit of Fig. 5.21 for the following case: $V_{DD} = -V_{SS} = 2.5$ V, $V_t = 1$ V, $\mu_n C_{ox} = 60\ \mu\text{A/V}^2$, $W/L = 120\ \mu\text{m}/3\ \mu\text{m}$, $I_D = 0.3$ mA, and $V_D = +0.4$ V.
Ans. $R_D = 7\text{ k}\Omega$; $R_S = 3.3\text{ k}\Omega$

Example 5.4

Figure 5.22 shows an NMOS transistor with its drain and gate terminals connected together. Find the i–v relationship of the resulting two-terminal device in terms of the MOSFET parameters $k_n = k_n'(W/L)$ and V_{tn}. Neglect channel-length modulation (i.e., $\lambda = 0$). Note that this two-terminal device is known as a **diode-connected transistor**.

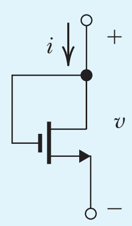


Figure 5.22

Solution

Since $v_D = v_G$ implies operation in the saturation mode,

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)(v_{GS} - V_{tn})^2$$

Now, $i = i_D$ and $v = v_{GS}$, thus

$$i = \frac{1}{2}k_n'\left(\frac{W}{L}\right)(v - V_{tn})^2$$

Replacing $k_n'\left(\frac{W}{L}\right)$ by k_n results in

$$i = \frac{1}{2}k_n(v - V_{tn})^2$$

EXERCISES

D5.9 For the circuit in Fig. E5.9, find the value of R that results in $V_D = 0.7$ V. The MOSFET has $V_{tn} = 0.5$ V, $\mu_n C_{ox} = 0.4$ mA/V^2, $W/L = \dfrac{0.72\ \mu\text{m}}{0.18\ \mu\text{m}}$, and $\lambda = 0$.
Ans. 34.4 kΩ

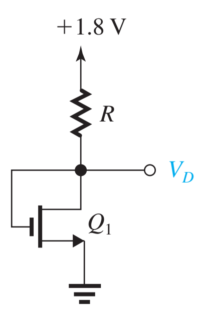


Figure E5.9

D5.10 Figure E5.10 shows a circuit obtained by augmenting the circuit of Fig. E5.9 considered in Exercise 5.9 with a transistor Q_2 identical to Q_1 and a resistance R_2. Find the value of R_2 that results in Q_2 operating at the edge of the saturation region. Use your solution to Exercise 5.9.
Ans. 50 kΩ

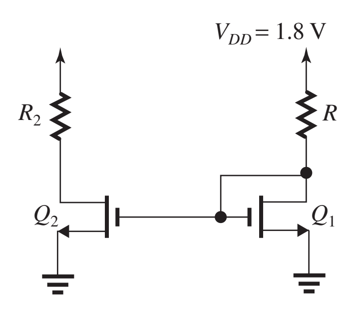


Figure E5.10

Example 5.5

Design the circuit in Fig. 5.23 to establish a drain voltage of 0.1 V. What is the effective resistance between drain and source at this operating point? Let $V_{tn} = 1$ V and $k_n'(W/L) = 1$ mA/ V^2.

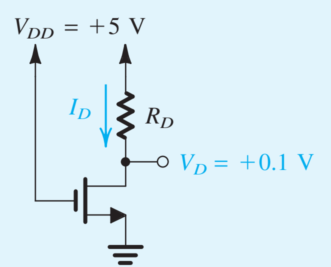


Figure 5.23 Circuit for Example 5.5.

Solution

Since the drain voltage is lower than the gate voltage by 4.9 V and $V_{tn} = 1$ V, the MOSFET is operating in the triode region. Thus the current I_D is given by

$$I_D = k_n' \frac{W}{L}\left[(V_{GS} - V_{tn})V_{DS} - \frac{1}{2}V_{DS}^2\right]$$

$$I_D = 1 \times \left[(5-1) \times 0.1 - \frac{1}{2} \times 0.01\right]$$

$$= 0.395 \text{ mA}$$

The required value for R_D can be found as follows:

$$R_D = \frac{V_{DD} - V_D}{I_D}$$
$$= \frac{5 - 0.1}{0.395} = 12.4\ \text{k}\Omega$$

In a practical discrete-circuit design problem, one selects the closest standard value available for, say, 5% resistors—in this case, 12 kΩ; see Appendix J. Since the transistor is operating in the triode region with a small V_{DS}, the effective drain-to-source resistance can be determined as follows:

$$r_{DS} = \frac{V_{DS}}{I_D}$$
$$= \frac{0.1}{0.395} = 253\ \Omega$$

Alternatively, we can determine r_{DS} by using the formula

$$r_{DS} = \frac{1}{k_n V_{OV}}$$

to obtain

$$r_{DS} = \frac{1}{1 \times (5-1)} = 0.25\ \text{k}\Omega = 250\ \Omega$$

which is close to the value found above.

EXERCISE

5.11 If in the circuit of Example 5.5 the value of R_D is doubled, find approximate values for I_D and V_D.
Ans. 0.2 mA; 0.05 V

Example 5.6

Analyze the circuit shown in Fig. 5.24(a) to determine the voltages at all nodes and the currents through all branches. Let $V_{tn} = 1$ V and $k_n'(W/L) = 1$ mA/V^2. Neglect the channel-length modulation effect (i.e., assume $\lambda = 0$).

Example 5.6 *continued*

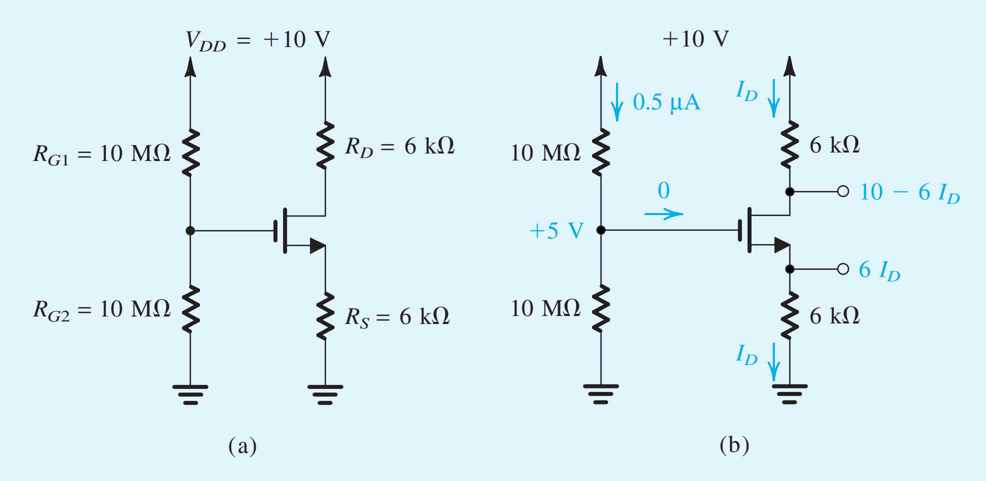


Figure 5.24 **(a)** Circuit for Example 5.6. **(b)** The circuit with some of the analysis details shown.

Solution

Since the gate current is zero, the voltage at the gate is simply determined by the voltage divider formed by the two 10-MΩ resistors,

$$V_G = V_{DD}\frac{R_{G2}}{R_{G2}+R_{G1}} = 10 \times \frac{10}{10+10} = +5\text{ V}$$

With this positive voltage at the gate, the NMOS transistor will be turned on. We do not know, however, whether the transistor will be operating in the saturation region or in the triode region. We shall assume saturation-region operation, solve the problem, and then check the validity of our assumption. Obviously, if our assumption turns out not to be valid, we will have to solve the problem again for triode-region operation.

Refer to Fig. 5.24(b). Since the voltage at the gate is 5 V and the voltage at the source is I_D (mA) × 6 (kΩ) = $6I_D$ (V), we have

$$V_{GS} = 5 - 6I_D$$

Thus I_D is given by

$$I_D = \frac{1}{2}k_n'\frac{W}{L}\left(V_{GS}-V_{tn}\right)^2$$

$$= \frac{1}{2} \times 1 \times \left(5 - 6I_D - 1\right)^2$$

which results in the following quadratic equation in I_D:

$$18I_D^2 - 25I_D + 8 = 0$$

This equation yields two values for I_D: 0.89 mA and 0.5 mA. The first value results in a source voltage of $6 \times 0.89 = 5.34$ V, which is greater than the gate voltage and does not make physical sense as it would imply that the NMOS transistor is cut off. Thus,

$$I_D = 0.5 \text{ mA}$$
$$V_S = 0.5 \times 6 = +3 \text{ V}$$
$$V_{GS} = 5 - 3 = 2 \text{ V}$$
$$V_D = 10 - 6 \times 0.5 = +7 \text{ V}$$

Since $V_D > V_G - V_{tn}$, the transistor is operating in saturation, as initially assumed.

EXERCISES

5.12 For the circuit of Fig. 5.24, what is the largest value that R_D can have while the transistor remains in the saturation mode?
Ans. 12 kΩ

D5.13 Redesign the circuit of Fig. 5.24 for the following requirements: $V_{DD} = +5$ V, $I_D = 0.32$ mA, $V_S = 1.6$ V, $V_D = 3.4$ V, with a 1-μA current through the voltage divider R_{G1}, R_{G2}. Assume the same MOSFET as in Example 5.6.
Ans. $R_{G1} = 1.6$ MΩ; $R_{G2} = 3.4$ MΩ, $R_S = R_D = 5$ kΩ

Example 5.7

Design the circuit of Fig. 5.25 so that the transistor operates in saturation with $I_D = 0.5$ mA and $V_D = +3$ V. Let the PMOS transistor have $V_{tp} = -1$ V and $k_p'(W/L) = 1$ mA/V^2. Assume $\lambda = 0$. What is the largest value that R_D can have while maintaining saturation-region operation?

Example 5.7 *continued*

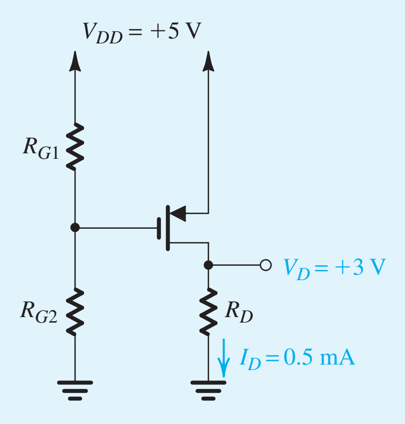


Figure 5.25 Circuit for Example 5.7.

Solution

Since the MOSFET is to be in saturation, we can write

$$I_D = \frac{1}{2} k_p' \frac{W}{L} |V_{OV}|^2$$

Substituting $I_D = 0.5$ mA and $k_p' W/L = 1$ mA/V^2, we obtain

$$|V_{OV}| = 1 \text{ V}$$

and

$$V_{SG} = |V_{tp}| + |V_{OV}| = 1 + 1 = 2 \text{ V}$$

Since the source is at +5 V, the gate voltage must be set to +3 V. This can be achieved by the appropriate selection of the values of R_{G1} and R_{G2}. A possible selection is $R_{G1} = 2$ MΩ and $R_{G2} = 3$ MΩ.

The value of R_D can be found from

$$R_D = \frac{V_D}{I_D} = \frac{3}{0.5} = 6 \text{ k}\Omega$$

Saturation-mode operation will be maintained up to the point that V_D exceeds V_G by $|V_{tp}|$; that is, until

$$V_{D\max} = 3 + 1 = 4 \text{ V}$$

This value of drain voltage is obtained with R_D given by

$$R_D = \frac{4}{0.5} = 8 \text{ k}\Omega$$

EXERCISE

D5.14 For the circuit in Fig. E5.14, find the value of R that results in the PMOS transistor operating with an overdrive voltage $|V_{OV}| = 0.6$ V. The threshold voltage is $V_{tp} = -0.4$ V, the process transconductance parameter $k_p' = 0.1$ mA/V^2, and $W/L = 10$ μm/0.18 μm.
Ans. 800 Ω

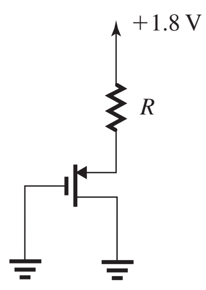


Figure E5.14

Example 5.8

The NMOS and PMOS transistors in the circuit of Fig. 5.26(a) are matched, with $k_n'(W_n/L_n) = k_p'(W_p/L_p) = 1$ mA/V^2 and $V_{tn} = -V_{tp} = 1$ V. Assuming $\lambda = 0$ for both devices, find the drain currents i_{DN} and i_{DP}, as well as the voltage v_O, for $v_I = 0$ V, +2.5 V, and −2.5 V.

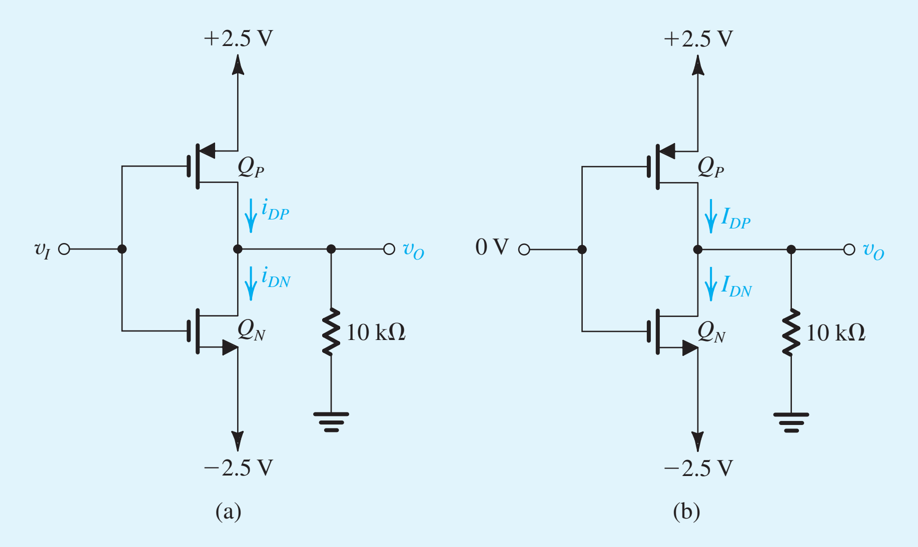


Figure 5.26 Circuits for Example 5.8.

Example 5.8 *continued*

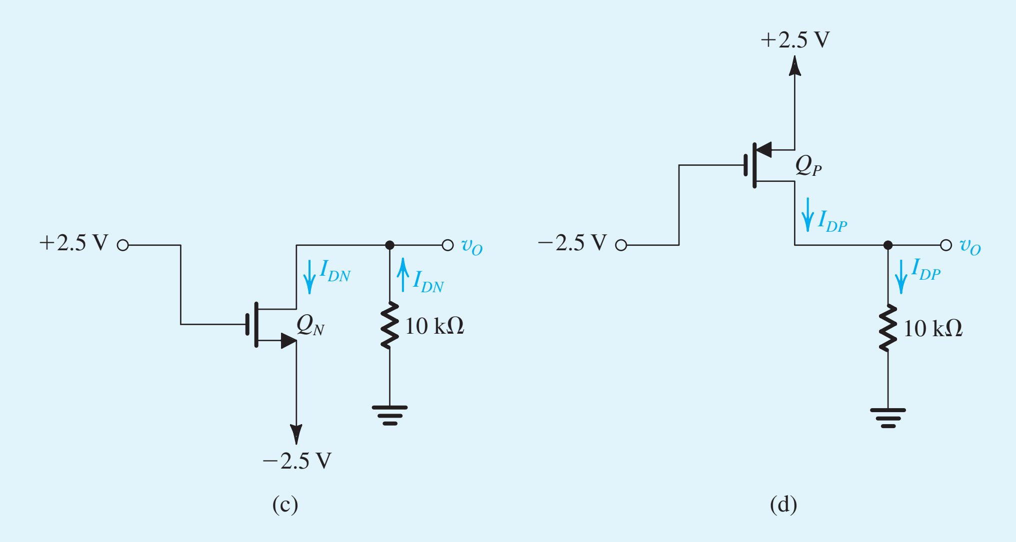


Figure 5.26 *continued*

Solution

Figure 5.26(b) shows the circuit for the case $v_I = 0$ V. We note that since Q_N and Q_P are perfectly matched and are operating at equal values of $|V_{GS}| = 2.5$ V, the circuit is symmetrical, which dictates that $v_O = 0$ V. Thus both Q_N and Q_P are operating with $|V_{DG}| = 0$ and, hence, in saturation. The drain currents can now be found from

$$I_{DP} = I_{DN} = \tfrac{1}{2} \times 1 \times (2.5 - 1)^2 = 1.125 \text{ mA}$$

Next, we consider the circuit with $v_I = +2.5$ V. Transistor Q_P will have a V_{SG} of zero and thus will be cut off, reducing the circuit to that shown in Fig. 5.26(c). We note that v_O will be negative, and thus v_{GD} will be greater than V_{tn}, causing Q_N to operate in the triode region. For simplicity we shall assume that v_{DS} is small and thus use

$$\begin{aligned} I_{DN} &\simeq k_n'(W_n/L_n)(V_{GS} - V_{tn})V_{DS} \\ &= 1[2.5 - (-2.5) - 1][v_O - (-2.5)] \end{aligned}$$

From the circuit diagram shown in Fig. 5.26(c), we can also write

$$I_{DN}(\text{mA}) = \frac{0 - v_O}{10(\text{k}\Omega)}$$

These two equations can be solved simultaneously to yield

$$I_{DN} = 0.244 \text{ mA} \qquad v_O = -2.44 \text{ V}$$

Note that $V_{DS} = -2.44 - (-2.5) = 0.06$ V, which is small as assumed.

Finally, the situation for the case $v_I = -2.5$ V [Fig. 5.26(d)] will be the exact complement of the case $v_I = +2.5$ V: Transistor Q_N will be off. Thus $I_{DN} = 0$, Q_P will be operating in the triode region with $I_{DP} = 0.244$ mA and $v_O = +2.44$ V.

EXERCISE

5.15 The NMOS and PMOS transistors in the circuit of Fig. E5.15 are matched with $k_n'(W_n/L_n) = k_p'(W_p/L_p) = 1$ mA/V^2 and $V_{tn} = -V_{tp} = 1$ V. Assuming $\lambda = 0$ for both devices, find the drain currents i_{DN} and i_{DP} and the voltage v_O for $v_I = 0$ V, +2.5 V, and –2.5 V.

Ans. $v_I = 0$ V: 0 mA, 0 mA, 0 V; $v_I = +2.5$ V: 0.104 mA, 0 mA, 1.04 V; $v_I = -2.5$ V: 0 mA, 0.104 mA, –1.04 V

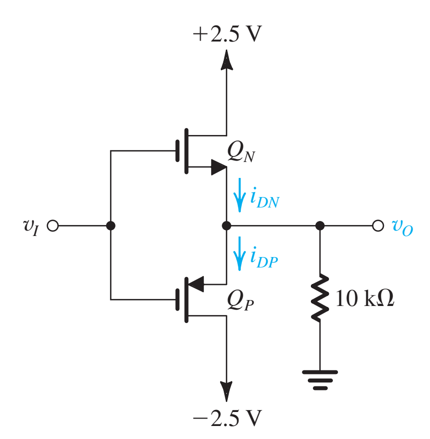


Figure E5.15

Concluding Remark If a MOSFET is conducting but its mode of operation (saturation or triode) is not known, we assume operation in the saturation region, solve the problem, and check whether the conditions for saturation-mode operation are satisfied. If not, then the MOSFET is operating in the triode region and the analysis is done accordingly.

GORDON MOORE—HIS LAW:

A half-century ago, Gordon Moore, who would go on to become a cofounder first of Fairchild Semiconductor and then of Intel, presented a startling idea in the issue of *Electronics Magazine* dated April 19, 1965. Moore, who had a doctorate in chemistry, had projected the potential growth of the integrated-circuit industry based on five points spanning a seven-year period from 1958 to 1965. The conclusion he reached—that the number of transistors per chip had been increasing and would continue to increase by a factor of 2 every two years or so—was destined to propel progress in integrated circuits over the succeeding decades into the twenty-first century. Doubling of the number of transistors was predicted on the basis of another prediction: the continuing shrinkage of transistor dimensions. In early recognition of the importance of this prediction, Carver Mead, a pioneer in very large scale integration (VLSI), soon began to refer to this prediction as "Moore's law." (See Chapter 15, Section 15.1, for the implications of Moore's law).

5.4 The Body Effect and Other Topics

In this section we briefly consider a number of important though secondary issues.

5.4.1 The Role of the Substrate—The Body Effect

In many applications the source terminal is connected to the substrate (or body) terminal B, which results in the *pn* junction between the substrate and the induced channel (review Fig. 5.5) having a constant zero (cutoff) bias. In such a case the substrate does not play any role in circuit operation and its existence can be ignored altogether.

In integrated circuits, however, the substrate is usually common to many MOS transistors. In order to maintain the cutoff condition for all the substrate-to-channel junctions, the substrate is usually connected to the most negative power supply in an NMOS circuit (the most positive in a PMOS circuit). The resulting reverse-bias voltage between source and body (V_{SB} in an n-channel device) will have an effect on device operation. To appreciate this fact, consider an NMOS transistor and let its substrate be made negative relative to the source. The reverse-bias voltage will widen the depletion region (refer to Fig. 5.2). This in turn reduces the channel depth. To return the channel to its former state, v_{GS} has to be increased.

The effect of V_{SB} on the channel can be most conveniently represented as a change in the threshold voltage V_t. Specifically, it has been shown that increasing the reverse substrate bias voltage V_{SB} results in an increase in V_t according to the relationship

$$V_t = V_{t0} + \gamma\left[\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f}\right] \tag{5.30}$$

where V_{t0} is the threshold voltage for $V_{SB} = 0$; ϕ_f is a physical parameter with $(2\phi_f)$ typically 0.6 V; γ is a fabrication-process parameter given by

$$\gamma = \frac{\sqrt{2qN_A\epsilon_s}}{C_{ox}} \tag{5.31}$$

where q is the magnitude of the electron charge (1.6×10^{-19} C), N_A is the doping concentration of the p-type substrate, and ϵ_s is the permittivity of silicon ($11.7\epsilon_0 = 11.7 \times 8.854 \times 10^{-14} = 1.04 \times 10^{-12}$ F/cm). The parameter γ has the dimension of $\sqrt{\text{V}}$ and is typically 0.4 $\text{V}^{1/2}$. Finally, note that Eq. (5.30) applies equally well for p-channel devices with V_{SB} replaced by

the reverse bias of the substrate, V_{BS} (or, alternatively, replace V_{SB} by $|V_{SB}|$) and note that γ is negative. Also, in evaluating γ, N_A must be replaced with N_D, the doping concentration of the n well in which the PMOS is formed. For p-channel devices, $2\phi_f$ is typically 0.75 V, and γ is typically –0.5 $V^{1/2}$.

EXERCISE

5.16 An NMOS transistor has $V_{t0} = 0.8$ V, $2\phi_f = 0.7$ V, and $\gamma = 0.4\ V^{1/2}$. Find V_t when $V_{SB} = 3$ V.
Ans. 1.23 V

Equation (5.30) indicates that an incremental change in V_{SB} gives rise to an incremental change in V_t, which in turn results in an incremental change in i_D even though v_{GS} might have been kept constant. It follows that the body voltage controls i_D; thus the body acts as another gate for the MOSFET, a phenomenon known as the **body effect**. Here we note that the parameter γ is known as the **body-effect parameter**.

5.4.2 Temperature Effects

Both V_t and k' are temperature sensitive. The magnitude of V_t decreases by about 2 mV for every 1°C rise in temperature. This decrease in $|V_t|$ gives rise to a corresponding increase in drain current as temperature is increased. However, because k' decreases with temperature and its effect is a dominant one, the overall observed effect of a temperature increase is a *decrease* in drain current. This very interesting result is put to use in applying the MOSFET in power circuits (Chapter 12).

5.4.3 Breakdown and Input Protection

As the voltage on the drain is increased, a value is reached at which the *pn* junction between the drain region and substrate suffers avalanche breakdown (see Section 3.5.3). This breakdown usually occurs at voltages of 20 V to 150 V and results in a somewhat rapid increase in current (known as a **weak avalanche**).

Another breakdown effect that occurs at lower voltages (about 20 V) in modern devices is called **punch-through**. It occurs in devices with relatively short channels when the drain voltage is increased to the point that the depletion region surrounding the drain region extends through the channel to the source. The drain current then increases rapidly. Normally, punch-through does not result in permanent damage to the device.

Yet another kind of breakdown occurs when the gate-to-source voltage exceeds about 30 V. This is the breakdown of the gate oxide and results in permanent damage to the device. Although 30 V may seem high, it must be remembered that the MOSFET has a very high input resistance and a very small input capacitance, and thus small amounts of static charge accumulating on the gate capacitor can cause its breakdown voltage to be exceeded.

To prevent the accumulation of static charge on the gate capacitor of a MOSFET, gate-protection devices are usually included at the input terminals of MOS integrated circuits. The protection mechanism invariably makes use of clamping diodes.

5.4.4 Velocity Saturation

At high longitudinal electric fields, the drift velocity of charge carriers in the channel reaches an upper limit (approximately 10^7 cm/s for electrons and holes in silicon). This effect, which in modern very-short-channel devices can occur for v_{DS} lower than 1 V, is called velocity saturation. It can be shown that when velocity saturation occurs, the current i_D will no longer be related to v_{GS} by the square-law relationship. Rather, i_D becomes linearly dependent on v_{GS} and the transconductance g_m becomes constant and independent of v_{GS}. In Chapter 15, we shall consider velocity saturation in our study of deep-submicron (i.e., $L < 0.25$ µm) CMOS digital circuits.

5.4.5 The Depletion-Type MOSFET

We conclude this section with a brief discussion of another type of MOSFET, the depletion-type MOSFET. Its structure is similar to that of the enhancement-type MOSFET with one important difference: The depletion MOSFET has a physically implanted channel. Thus an n-channel depletion-type MOSFET has an n-type silicon region connecting the n^+ source and the n^+ drain regions at the top of the p-type substrate. Thus if a voltage v_{DS} is applied between drain and source, a current i_D flows for $v_{GS} = 0$. In other words, there is no need to induce a channel, unlike the case of the enhancement MOSFET.

The channel depth and hence its conductivity can be controlled by v_{GS} in exactly the same manner as in the enhancement-type device. Applying a positive v_{GS} enhances the channel by attracting more electrons into it. Here, however, we also can apply a negative v_{GS}, which causes electrons to be repelled from the channel, and thus the channel becomes shallower and its conductivity decreases. The negative v_{GS} is said to **deplete** the channel of its charge carriers,

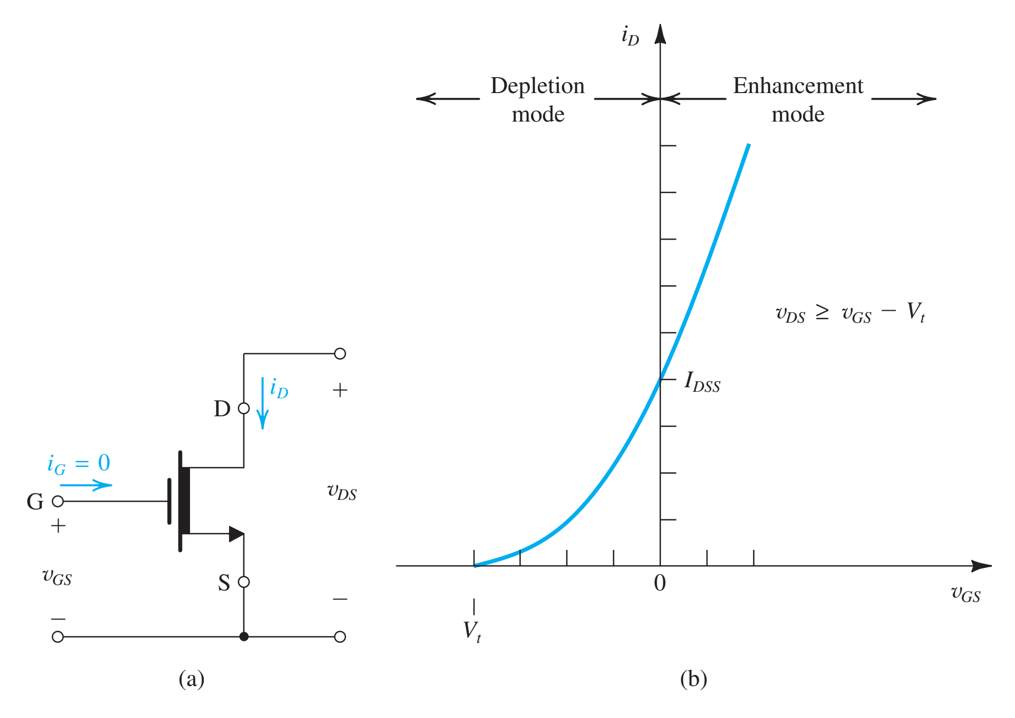


Figure 5.27 The circuit symbol **(a)** and the i_D–v_{GS} characteristic in saturation **(b)** for an n-channel depletion-type MOSFET.

and this mode of operation (negative v_{GS}) is called **depletion mode**. As the magnitude of v_{GS} is increased in the negative direction, a value is reached at which the channel is completely depleted of charge carriers and i_D is reduced to zero even though v_{DS} may be still applied. This negative value of v_{GS} is the threshold voltage of the n-channel depletion-type MOSFET.

The description above suggests (correctly) that a depletion-type MOSFET can be operated in the enhancement mode by applying a positive v_{GS} and in the depletion mode by applying a negative v_{GS}. This is illustrated in Fig. 5.27, which shows both the circuit symbol for the depletion NMOS transistor (Fig. 5.27a) and its i_D–v_{GS} characteristic. Observe that here the threshold voltage V_{tn} is negative. The i_D–v_{DS} characteristics (not shown) are similar to those for the enhancement-type MOSFET except for the negative V_{tn}. Finally, note that the device symbol denotes the existing channel via the shaded area next to the vertical line.

Depletion-type MOSFETs can be fabricated on the same IC chip as enhancement-type devices, resulting in circuits with improved characteristics, as will be shown in a later chapter. The depletion-type MOSFET, however, is a specialty device and is not commonly used.

EXERCISE

5.17 For a depletion-type NMOS transistor with $V_t = -2$ V and $k_n'(W/L) = 2$ mA/V^2, find the minimum v_{DS} required to operate in the saturation region when $v_{GS} = +1$ V. What is the corresponding value of i_D?
Ans. 3 V; 9 mA

Summary

- The enhancement-type MOSFET is currently the most widely used semiconductor device. It is the basis of CMOS technology, which is the most popular IC fabrication technology at this time. CMOS provides both n-channel (NMOS) and p-channel (PMOS) transistors, which increases design flexibility. The minimum MOSFET channel length achievable with a given CMOS process is used to characterize the process. This figure has been continually reduced and is currently 32 nm.
- The overdrive voltage, $|v_{OV}| \equiv |v_{GS}| - |V_t|$, is the key quantity that governs the operation of the MOSFET. For the MOSFET to operate in the saturation region, which is the region for amplifier application, $|v_{DS}| \geq |v_{OV}|$, and the resulting $i_D = \frac{1}{2}\mu_n C_{ox}(W/L)v_{OV}^2$ (for NMOS; replace μ_n with μ_p for PMOS). If $|v_{DS}| < |v_{OV}|$, the MOSFET operates in the triode region, which together with cutoff is used for operating the MOSFET as a switch.
- Tables 5.1 and 5.2 provide summaries of the conditions and relationships that describe the operation of NMOS and PMOS transistors, respectively.
- In saturation, i_D shows some linear dependence on v_{DS} as a result of the change in channel length. This channel-length modulation phenomenon becomes more pronounced as L decreases. It is modeled by ascribing an output resistance $r_o = |V_A|/I_D$ to the MOSFET model. Here, the Early voltage $|V_A| = |V_A'|L$, where $|V_A'|$ is a process-dependent parameter.
- In the analysis of dc MOSFET circuits, if a MOSFET is conducting, but its region of operation (saturation or triode) is not known, one assumes saturation-mode operation. Then, one solves the problem and checks to determine whether the assumption was justified. If not, then the transistor is operating in the triode region, and the analysis is done accordingly.
- The depletion-type MOSFET has an implanted channel and thus can be operated in either the depletion or enhancement mode. It is characterized by the same equations used for the enhancement device except for having a negative V_{tn} (positive V_{tp} for depletion PMOS transistors).

Computer Simulations Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as allowable signal swing and amplifier nonlinear distortion. Instructions to assist in setting up PSPice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 5.1: Device Structure and Physical Operation

5.1 MOS technology is used to fabricate a capacitor, utilizing the gate metallization and the substrate as the capacitor electrodes. Find the area required per 1-pF capacitance for oxide thickness ranging from 2 nm to 10 nm. For a square plate capacitor of 10 pF, what dimensions are needed?

5.2 Calculate the total charge stored in the channel of an NMOS transistor having $C_{ox} = 9$ fF/μm^2, $L = 0.36$ μm, and $W = 3.6$ μm, and operated at $V_{OV} = 0.2$ V and $V_{DS} = 0$ V.

5.3 Use dimensional analysis to show that the units of the process transconductance parameter k_n' are A/V^2. What are the dimensions of the MOSFET transconductance parameter k_n?

5.4 An NMOS transistor that is operated with a small v_{DS} is found to exhibit a resistance r_{DS}. By what factor will r_{DS} change in each of the following situations?

(a) V_{OV} is doubled.
(b) The device is replaced with another fabricated in the same technology but with double the width.
(c) The device is replaced with another fabricated in the same technology but with both the width and length doubled.
(d) The device is replaced with another fabricated in a more advanced technology for which the oxide thickness is halved and similarly for W and L (assume μ_n remains unchanged).

D 5.5 An NMOS transistor fabricated in a technology for which $k_n' = 400$ μA/V^2 and $V_t = 0.5$ V is required to operate with a small v_{DS} as a variable resistor ranging in value from 250 Ω to 1 kΩ. Specify the range required for the control voltage V_{GS} and the required transistor width W. It is required to use the smallest possible device, as limited by the minimum channel length of this technology ($L_{min} = 0.18$ μm) and the maximum allowed voltage of 1.8 V.

5.6 Sketch a set of $i_D - v_{DS}$ characteristic curves for an NMOS transistor operating with a small v_{DS} (in the manner shown in Fig. 5.4). Let the MOSFET have $k_n = 5$ mA/V^2 and $V_{tn} = 0.5$ V. Sketch and clearly label the graphs for $V_{GS} = 0.5$, 1.0, 1.5, 2.0, and 2.5 V. Let V_{DS} be in the range 0 to 50 mV. Give the value of r_{DS} obtained for each of the five values of V_{GS}. Although only a sketch, your diagram should be drawn to scale as much as possible.

D 5.7 An n-channel MOS device in a technology for which oxide thickness is 4 nm, minimum channel length is 0.18 μm, $k_n' = 400$ μA/V^2, and $V_t = 0.5$ V operates in the triode region, with small v_{DS} and with the gate–source voltage in the range 0 V to +1.8 V. What device width is needed to ensure that the minimum available resistance is 1 kΩ?

5.8 Consider an NMOS transistor operating in the triode region with an overdrive voltage V_{OV}. Find an expression for the incremental resistance

$$r_{ds} \equiv 1 \Big/ \left. \frac{\partial i_D}{\partial v_{DS}} \right|_{v_{DS}=V_{DS}}$$

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

Give the values of r_{ds} in terms of k_n and V_{OV} for $V_{DS} = 0$, $0.2V_{OV}$, $0.5V_{OV}$, $0.8V_{OV}$, and V_{OV}.

5.9 An NMOS transistor with $k_n = 4\ \text{mA/V}^2$ and $V_t = 0.5$ V is operated with $V_{GS} = 1.0$ V. At what value of V_{DS} does the transistor enter the saturation region? What value of I_D is obtained in saturation?

5.10 Consider a CMOS process for which $L_{\min} = 0.25\ \mu\text{m}$, $t_{ox} = 6$ nm, $\mu_n = 460\ \text{cm}^2/\text{V}\cdot\text{s}$, and $V_t = 0.5$ V.

(a) Find C_{ox} and k_n'.
(b) For an NMOS transistor with $W/L = 20\ \mu\text{m}/0.25\ \mu\text{m}$, calculate the values of V_{OV}, V_{GS}, and $V_{DS\min}$ needed to operate the transistor in the saturation region with a dc current $I_D = 0.5$ mA.
(c) For the device in (b), find the values of V_{OV} and V_{GS} required to cause the device to operate as a 100-Ω resistor for very small v_{DS}.

5.11 A p-channel MOSFET with a threshold voltage $V_{tp} = -0.7$ V has its source connected to ground.

(a) What should the gate voltage be for the device to operate with an overdrive voltage of $|V_{OV}| = 0.4$ V?
(b) With the gate voltage as in (a), what is the highest voltage allowed at the drain while the device operates in the saturation region?
(c) If the drain current obtained in (b) is 0.5 mA, what would the current be for $V_D = -20$ mV and for $V_D = -2$V?

5.12 With the knowledge that $\mu_p = 0.4\mu_n$, what must be the relative width of n-channel and p-channel devices having equal channel lengths if they are to have equal drain currents when operated in the saturation mode with overdrive voltages of the same magnitude?

5.13 An n-channel device has $k_n' = 100\ \mu\text{A/V}^2$, $V_t = 0.7$ V, and $W/L = 20$. The device is to operate as a switch for small v_{DS}, utilizing a control voltage v_{GS} in the range 0 V to 5 V. Find the switch closure resistance, r_{DS}, and closure voltage, V_{DS}, obtained when $v_{GS} = 5$ V and $i_D = 1$ mA. If $\mu_p \simeq 0.4\mu_n$, what must W/L be for a p-channel device that provides the same performance as the n-channel device in this application?

5.14 Consider an n-channel MOSFET with $t_{ox} = 6$ nm, $\mu_n = 460\ \text{cm}^2/\text{V}\cdot\text{s}$, $V_t = 0.5$ V, and $W/L = 10$. Find the drain current in the following cases:

(a) $v_{GS} = 2.5$ V and $v_{DS} = 1$ V
(b) $v_{GS} = 2$ V and $v_{DS} = 1.5$ V
(c) $v_{GS} = 2.5$ V and $v_{DS} = 0.2$ V
(d) $v_{GS} = v_{DS} = 2.5$ V

***5.15** This problem illustrates the central point in the electronics revolution that has been in effect for the past four decades: By continually reducing the MOSFET size, we are able to pack more devices on an IC chip. Gordon Moore, co-founder of Intel Corporation, predicted this exponential growth of chip-packing density very early in the history of the development of the integrated circuit in the formulation that has become known as **Moore's law**.

The table on the next page shows four technology generations, each characterized by the minimum possible MOSFET channel length (row 1). In going from one generation to another, both L and t_{ox} are scaled by the same factor. The power supply utilized V_{DD} is also scaled by the same factor, to keep the magnitudes of all electrical fields within the device unchanged. Unfortunately, but for good reasons, V_t cannot be scaled similarly.

Complete the table entries, noting that row 5 asks for the transconductance parameter of an NMOS transistor with $W/L = 10$; row 9 asks for the value of I_D obtained with $V_{GS} = V_{DS} = V_{DD}$; row 10 asks for the power $P = V_{DD}I_D$ dissipated in the circuit. An important quantity is the power density, P/A, asked for in row 11. Finally, you are asked to find the number of transistors that can be placed on an IC chip fabricated in each of the technologies in terms of the number obtained with the 0.5-μm technology (n).

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

1	L (μm)	0.5	0.25	0.18	0.13
2	t_{ox} (nm)	10			
3	C_{ox} (fF/μm^2)				
4	k_n' (μA/V^2) ($\mu_n = 500$ cm^2/ V · s)				
5	k_n (mA/V^2) For $W/L = 10$				
6	Device area, A (μm^2)				
7	V_{DD} (V)	5			
8	V_t (V)	0.7	0.5	0.4	0.4
9	I_D (mA) For $V_{GS} = V_{DS} = V_{DD}$				
10	P (mW)				
11	P/A (mW/μm^2)				
12	Devices per chip	n			

Section 5.2: Current–Voltage Characteristics

In the following problems, when λ is not specified, assume it is zero.

5.16 Show that when channel-length modulation is neglected (i.e., $\lambda = 0$), plotting i_D/k_n versus v_{DS} for various values of v_{OV}, and plotting i_D/k_n versus v_{OV} for $v_{DS} \geq v_{OV}$, results in universal representation of the $i_D - v_{DS}$ and $i_D - v_{GS}$ characteristics of the NMOS transistor. That is, the resulting graphs are both technology and device independent. Furthermore, these graphs apply equally well to the PMOS transistor by a simple relabeling of variables. (How?) What is the slope at $v_{DS} = 0$ of each of the i_D/k_n versus v_{DS} graphs? For the i_D/k_n versus v_{OV} graph, find the slope at a point $v_{OV} = V_{OV}$.

5.17 An NMOS transistor having $V_t = 0.8$ V is operated in the triode region with v_{DS} small. With $V_{GS} = 1.2$ V, it is found to have a resistance r_{DS} of 1 kΩ. What value of V_{GS} is required to obtain $r_{DS} = 200$ Ω? Find the corresponding resistance values obtained with a device having twice the value of W.

5.18 A particular MOSFET for which $V_{tn} = 0.5$ V and $k_n'(W/L) = 1.6$ mA/V^2 is to be operated in the saturation region. If i_D is to be 50 μA, find the required v_{GS} and the minimum required v_{DS}. Repeat for $i_D = 200$ μA.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

5.19 A particular n-channel MOSFET is measured to have a drain current of 0.4 mA at $V_{GS} = V_{DS} = 1$ V and of 0.1 mA at $V_{GS} = V_{DS} = 0.8$ V. What are the values of k_n and V_t for this device?

D 5.20 For a particular IC-fabrication process, the transconductance parameter $k_n' = 400\ \mu\text{A/V}^2$, and $V_t = 0.5$ V. In an application in which $v_{GS} = v_{DS} = V_{\text{supply}} = 1.8$ V, a drain current of 2 mA is required of a device of minimum length of 0.18 μm. What value of channel width must the design use?

5.21 An NMOS transistor, operating in the linear-resistance region with $v_{DS} = 50$ mV, is found to conduct 25 μA for $v_{GS} = 1$ V and 50 μA for $v_{GS} = 1.5$ V. What is the apparent value of threshold voltage V_t? If $k_n' = 50\ \mu\text{A/V}^2$, what is the device W/L ratio? What current would you expect to flow with $v_{GS} = 2$ V and $v_{DS} = 0.1$ V? If the device is operated at $v_{GS} = 2$ V, at what value of v_{DS} will the drain end of the MOSFET channel just reach pinch-off, and what is the corresponding drain current?

5.22 For an NMOS transistor, for which $V_t = 0.4$ V, operating with v_{GS} in the range of 1.0 V to 1.8 V, what is the largest value of v_{DS} for which the channel remains continuous?

5.23 An NMOS transistor, fabricated with $W = 20$ μm and $L = 1$ μm in a technology for which $k_n' = 100\ \mu\text{A/V}^2$ and $V_t = 0.8$ V, is to be operated at very low values of v_{DS} as a linear resistor. For v_{GS} varying from 1.0 V to 4.8 V, what range of resistor values can be obtained? What is the available range if

(a) the device width is halved?
(b) the device length is halved?
(c) both the width and length are halved?

5.24 When the drain and gate of a MOSFET are connected together, a two-terminal device known as a "diode-connected transistor" results. Figure P5.24 shows such devices obtained from MOS transistors of both polarities. Show that

(a) the i–v relationship is given by

$$i = \frac{1}{2}k'\frac{W}{L}\left(v - |V_t|\right)^2$$

(b) the incremental resistance r for a device biased to operate at $v = |V_t| + V_{OV}$ is given by

$$r \equiv 1\Big/\left[\frac{\partial i}{\partial v}\right] = 1\Big/\left(k'\frac{W}{L}V_{OV}\right)$$

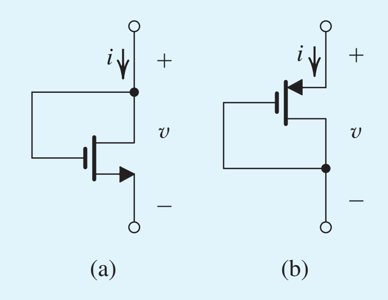


Figure P5.24

5.25 For the circuit in Fig. P5.25, sketch i_D versus v_S for v_S varying from 0 to V_{DD}. Clearly label your sketch.

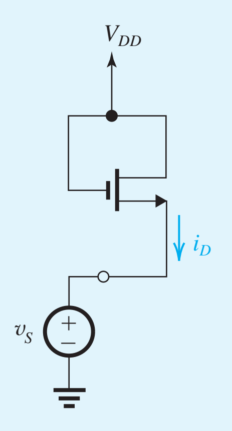


Figure P5.25

5.26 For the circuit in Fig. P5.26, find an expression for v_{DS} in terms of i_D. Sketch and clearly label a graph for v_{DS} versus i_D.

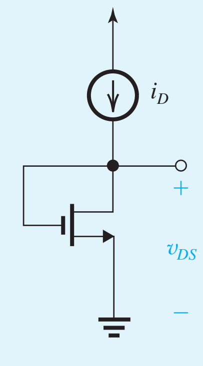


Figure P5.26

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

Case	Voltage (V)						Region of operation
	V_S	V_G	V_D	V_{GS}	V_{OV}	V_{DS}	
a	+1.0	+1.0	+2.0				
b	+1.0	+2.5	+2.0				
c	+1.0	+2.5	+1.5				
d	+1.0	+1.5	0				
e	0	+2.5	+1.0				
f	+1.0	+1.0	+1.0				
g	−1.0	0	0				
h	−1.5	0	0				
i	−1.0	0	+1.0				
j	+0.5	+2.0	+0.5				

***5.27** The table above lists 10 different cases labeled (a) to (j) for operating an NMOS transistor with $V_t = 1$ V. In each case the voltages at the source, gate, and drain (relative to the circuit ground) are specified. You are required to complete the table entries. Note that if you encounter a case for which v_{DS} is negative, you should exchange the drain and source before solving the problem. You can do this because the MOSFET is a symmetric device.

5.28 The NMOS transistor in Fig. P5.28 has $V_t = 0.4$ V and $k_n'(W/L) = 1$ mA/V^2. Sketch and clearly label i_D versus v_G with v_G varying in the range 0 to +1.8 V. Give equations for the various portions of the resulting graph.

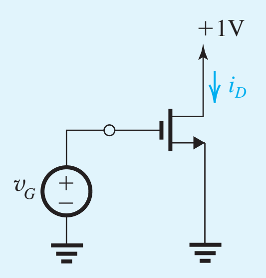


Figure P5.28

5.29 Figure P5.29 shows two NMOS transistors operating in saturation at equal V_{GS} and V_{DS}.

(a) If the two devices are matched except for a maximum possible mismatch in their W/L ratios of 3%, what is the maximum resulting mismatch in the drain currents?

(b) If the two devices are matched except for a maximum possible mismatch in their V_t values of 10 mV, what is the maximum resulting mismatch in the drain currents? Assume that the nominal value of V_t is 0.6 V.

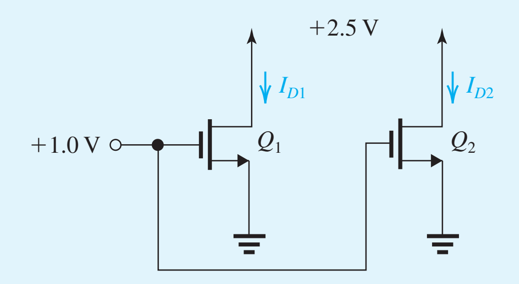


Figure P5.29

5.30 For a particular MOSFET operating in the saturation region at a constant v_{GS}, i_D is found to be 0.5 mA for $v_{DS} = 1$ V and 0.52 mA for $v_{DS} = 2$ V. What values of r_o, V_A, and λ correspond?

5.31 A particular MOSFET has $V_A = 20$ V. For operation at 0.1 mA and 1 mA, what are the expected output resistances? In each case, for a change in v_{DS} of 1 V, what percentage change in drain current would you expect?

D 5.32 In a particular IC design in which the standard channel length is 1 μm, an NMOS device with W/L of 10 operating at 200 μA is found to have an output resistance of 100 kΩ, about $\frac{1}{5}$ of that needed. What dimensional change can be made to solve the problem? What is the new device length? The new device width? The new W/L ratio? What is V_A for the standard device in this IC? The new device?

D 5.33 For a particular n-channel MOS technology, in which the minimum channel length is 0.5 μm, the associated value of λ is 0.03 V^{-1}. If a particular device for which L is 1.5 μm operates in saturation at $v_{DS} = 1$ V with a drain current of 100 μA, what does the drain current become if v_{DS} is raised to 5 V? What percentage change does this represent? What can be done to reduce the percentage by a factor of 2?

5.34 An NMOS transistor is fabricated in a 0.5-μm process having $k_n' = 200$ μA/V^2 and $V_A' = 20$ V/μm of channel length. If $L = 1.5$ μm and $W = 15$ μm, find V_A and λ. Find the value of I_D that results when the device is operated with an overdrive voltage of 0.5 V and $V_{DS} = 2$ V. Also, find the value of r_o at this operating point. If V_{DS} is increased by 1 V, what is the corresponding change in I_D?

5.35 If in an NMOS transistor, both W and L are quadrupled and V_{OV} is halved, by what factor does r_o change?

D 5.36 Consider the circuit in Fig. P5.29 with both transistors perfectly matched but with the dc voltage at the drain of Q_1 lowered to +2 V. If the two drain currents are to be matched within 1% (i.e., the maximum difference allowed between the two currents is 1%), what is the minimum required value of V_A? If the technology is specified to have $V_A' = 100$ V/μm, what is the minimum channel length the designer must use?

5.37 Complete the missing entries in the following table, which describes characteristics of suitably biased NMOS transistors:

MOS	1	2	3	4
λ (V^{-1})		0.02		
V_A (V)	20			100
I_D (mA)	0.5		0.1	
r_o (kΩ)		25	100	500

5.38 A PMOS transistor has $k_p'(W/L) = 100$ μA/V^2, $V_t = -1.0$ V, and $\lambda = -0.02$ V^{-1}. The gate is connected to ground and the source to +5 V. Find the drain current for $v_D = +4$ V, +2 V, +1 V, 0 V, and –5 V.

5.39 A p-channel transistor for which $|V_t| = 0.8$ V and $|V_A| = 40$ V operates in saturation with $|v_{GS}| = 3$ V, $|v_{DS}| = 4$ V, and $i_D = 3$ mA. Find corresponding signed values for v_{GS}, v_{SG}, v_{DS}, v_{SD}, V_t, V_A, λ, and $k_p'(W/L)$.

5.40 The table below lists the terminal voltages of a PMOS transistor in six cases, labeled a, b, c, d, e, and f. The transistor has $V_{tp} = -1$ V. Complete the table entries.

	V_S	V_G	V_D	V_{SG}	$\lvert V_{OV}\rvert$	V_{SD}	Region of operation
a	+2	+2	0				
b	+2	+1	0				
c	+2	0	0				
d	+2	0	+1				
e	+2	0	+1.5				
f	+2	0	+2				

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

5.41 The PMOS transistor in Fig. P5.41 has $V_{tp} = -0.5$ V. As the gate voltage v_G is varied from +3 V to 0 V, the transistor moves through all of its three possible modes of operation. Specify the values of v_G at which the device changes modes of operation.

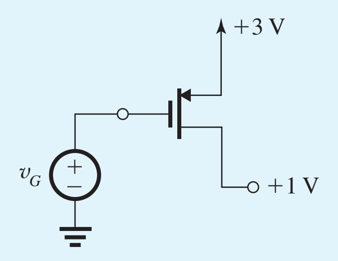


Figure P5.41

***5.42** Various NMOS and PMOS transistors, numbered 1 to 4, are measured in operation, as shown in the table at the bottom of the page. For each transistor, find the values of $\mu C_{ox}W/L$ and V_t that apply and complete the table, with V in volts, I in μA, and $\mu C_{ox}W/L$ in μA/V². Assume $\lambda = 0$.

***5.43** All the transistors in the circuits shown in Fig. P5.43 have the same values of $|V_t|$, k', W/L, and λ. Moreover, λ is negligibly small. All operate in saturation at $I_D = I$ and $|V_{GS}| = |V_{DS}| = 1$ V. Find the voltages V_1, V_2, V_3, and V_4. If $|V_t| = 0.5$ V and $I = 0.1$ mA, how large a resistor can be inserted in series with each drain while maintaining saturation? If the current source I requires at least 0.5 V between its terminals to operate properly, what is the largest resistor that can be placed in series with each MOSFET source while ensuring saturated-mode operation of each transistor at $I_D = I$? In the latter limiting situation, what do V_1, V_2, V_3, and V_4 become?

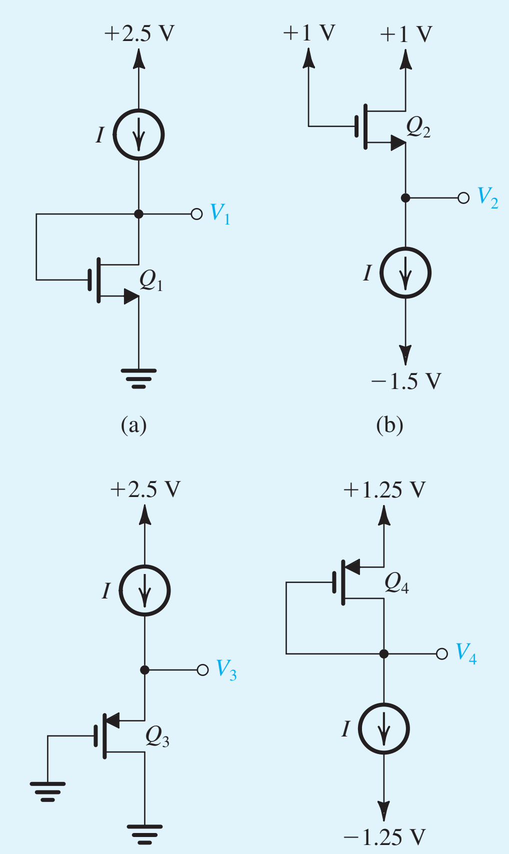


Figure P5.43

Case	Transistor	V_S	V_G	V_D	I_D	Type	Mode	$\mu C_{ox}W/L$	V_t
a	1	0	1	2.5	100				
	1	0	1.5	2.5	400				
b	2	5	3	−4.5	50				
	2	5	2	−0.5	450				
c	3	5	3	4	200				
	3	5	2	0	800				
d	4	−2	0	0	72				
	4	−4	0	−3	270				

Section 5.3: MOSFET Circuits at DC

Note: If λ is not specified, assume it is zero.

D 5.44 Design the circuit of Fig. P5.44 to establish a drain current of 0.1 mA and a drain voltage of +0.3 V. The MOSFET has $V_t = 0.5$ V, $\mu_n C_{ox} = 400\ \mu\text{A/V}^2$, $L = 0.4\ \mu$m, and $W = 5\ \mu$m. Specify the required values for R_S and R_D.

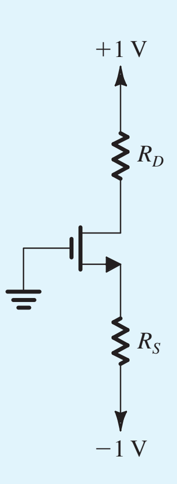


Figure P5.44

5.45 The NMOS transistor in the circuit of Fig. P5.44 has $V_t = 0.4$ V and $k_n = 4\ \text{mA/V}^2$. The voltages at the source and the drain are measured and found to be −0.6 V and +0.2 V, respectively. What current I_D is flowing, and what must the values of R_D and R_S be? What is the largest value for R_D for which I_D remains unchanged from the value found?

D 5.46 For the circuit in Fig. E5.10, assume that Q_1 and Q_2 are matched except for having different widths, W_1 and W_2. Let $V_t = 0.5$ V, $k_n' = 0.4\ \text{mA/V}^2$, $L_1 = L_2 = 0.36\ \mu$m, $W_1 = 1.44\ \mu$m, and $\lambda = 0$.

(a) Find the value of R required to establish a current of 50 μA in Q_1.

(b) Find W_2 and R_2 so that Q_2 operates at the edge of saturation with a current of 0.5 mA.

5.47 The transistor in the circuit of Fig. P5.47 has $k_n' = 0.4\ \text{mA/V}^2$, $V_t = 0.4$ V, and $\lambda = 0$. Show that operation at the edge of saturation is obtained when the following condition is satisfied:

$$\left(\frac{W}{L}\right)R_D \simeq 2.5\ \text{k}\Omega$$

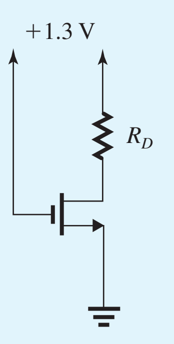


Figure P5.47

D 5.48 It is required to operate the transistor in the circuit of Fig. P5.47 at the edge of saturation with $I_D = 0.1$ mA. If $V_t = 0.4$ V, find the required value of R_D.

D 5.49 The PMOS transistor in the circuit of Fig. P5.49 has $V_t = -0.5$ V, $\mu_p C_{ox} = 100\ \mu\text{A/V}^2$, $L = 0.18\ \mu$m, and $\lambda = 0$. Find the values required for W and R in order to establish a drain current of 180 μA and a voltage V_D of 1 V.

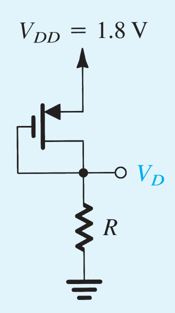


Figure P5.49

D 5.50 The NMOS transistors in the circuit of Fig. P5.50 have $V_t = 0.5$ V, $\mu_n C_{ox} = 250\ \mu\text{A/V}^2$, $\lambda = 0$, and $L_1 = L_2 = 0.25\ \mu$m. Find the required values of gate width for each of Q_1

and Q_2, and the value of R, to obtain the voltage and current values indicated.

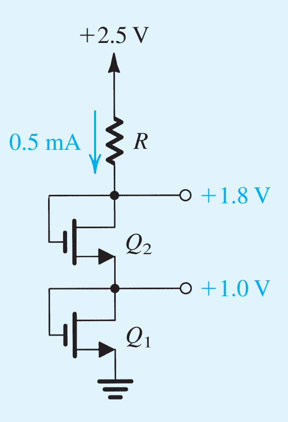


Figure P5.50

D 5.51 The NMOS transistors in the circuit of Fig. P5.51 have $V_t = 0.5$ V, $\mu_n C_{ox} = 90\ \mu\text{A/V}^2$, $\lambda = 0$, and $L_1 = L_2 = L_3 = 0.5\ \mu$m. Find the required values of gate width for each of Q_1, Q_2, and Q_3 to obtain the voltage and current values indicated.

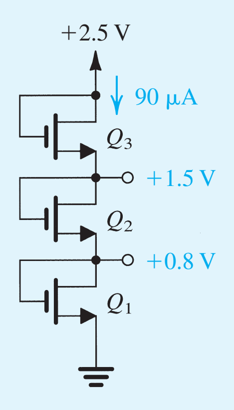


Figure P5.51

5.52 Consider the circuit of Fig. 5.24(a). In Example 5.5 it was found that when $V_t = 1$ V and $k_n'(W/L) = 1\ \text{mA/V}^2$, the drain current is 0.5 mA and the drain voltage is +7 V. If the transistor is replaced with another having $V_t = 1.5$ V with $k_n'(W/L) = 1.5\ \text{mA/V}^2$, find the new values of I_D and V_D. Comment on how tolerant (or intolerant) the circuit is to changes in device parameters.

D 5.53 Using a PMOS transistor with $V_t = -1.5$ V, $k_p'(W/L) = 4\ \text{mA/V}^2$, and $\lambda = 0$, design a circuit that resembles that in Fig. 5.24(a). Using a 10-V supply, design for a gate voltage of +6 V, a drain current of 0.5 mA, and a drain voltage of +5 V. Find the values of R_S and R_D. Also, find the values of the resistances in the voltage divider feeding the gate, assuming a 1-µA current in the divider.

5.54 The MOSFET in Fig. P5.54 has $V_t = 0.4$ V, $k_n' = 500\ \mu\text{A/V}^2$, and $\lambda = 0$. Find the required values of W/L and of R so that when $v_I = V_{DD} = +1.3$ V, $r_{DS} = 50\ \Omega$ and $v_O = 50$ mV.

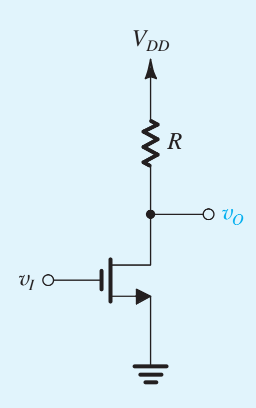


Figure P5.54

5.55 In the circuits shown in Fig. P5.55, transistors are characterized by $|V_t| = 1$ V, $k'W/L = 4\ \text{mA/V}^2$, and $\lambda = 0$.

(a) Find the labeled voltages V_1 through V_7.

(b) In each of the circuits, replace the current source with a resistor. Select the resistor value to yield a current as close to that of the current source as possible, while using resistors specified in the 1% table provided in Appendix J.

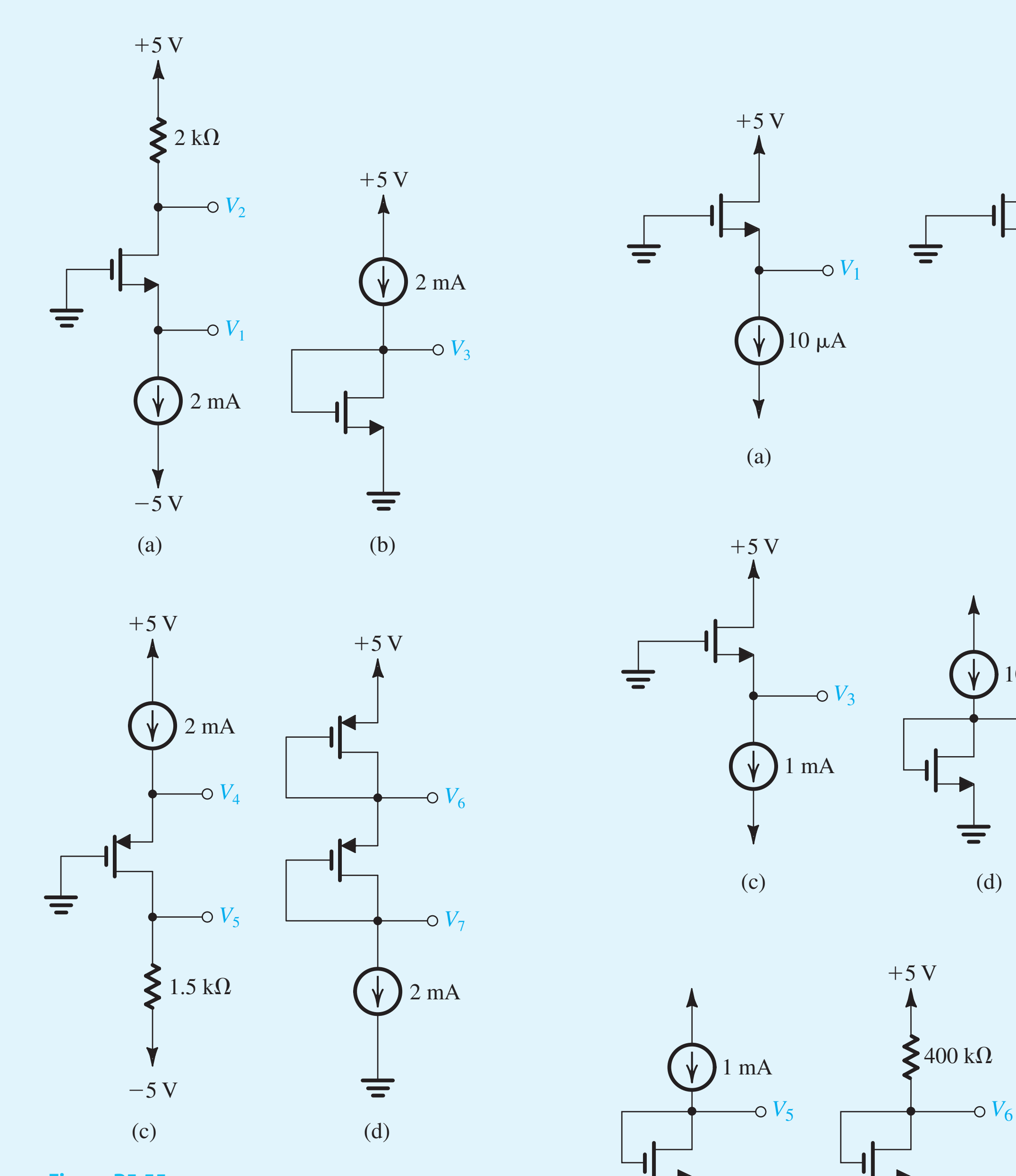


Figure P5.55

5.56 For each of the circuits in Fig. P5.56, find the labeled node voltages. For all transistors, $k_n'(W/L) = 0.5\ \text{mA/V}^2$, $V_t = 0.8$ V, and $\lambda = 0$.

Figure P5.56 *continued*

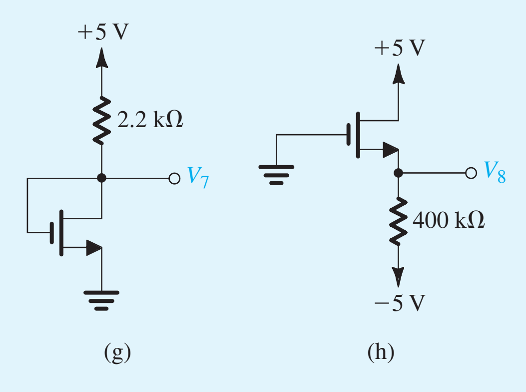


Figure P5.56 *continued*

5.57 For each of the circuits shown in Fig. P5.57, find the labeled node voltages. The NMOS transistors have $V_t = 0.9$ V and $k_n'(W/L) = 1.5$ mA/V^2.

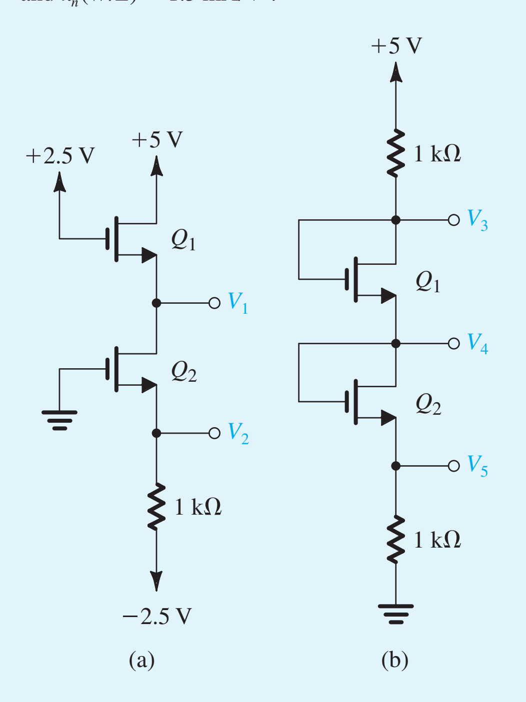


Figure P5.57

***5.58** For the circuit in Fig. P5.58:

(a) Show that for the PMOS transistor to operate in saturation, the following condition must be satisfied:

$$IR \le |V_{tp}|$$

(b) If the transistor is specified to have $|V_{tp}| = 1$ V and $k_p = 0.2$ mA/V^2, and for $I = 0.1$ mA, find the voltages V_{SD} and V_{SG} for $R = 0$, 10 kΩ, 30 kΩ, and 100 kΩ.

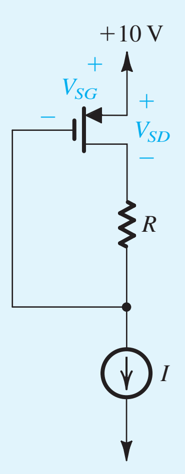


Figure P5.58

5.59 For the circuits in Fig. P5.59, $\mu_n C_{ox} = 3\mu_p C_{ox} = 270\ \mu\text{A/V}^2$, $|V_t| = 0.5$ V, $\lambda = 0$, $L = 1\ \mu$m, and $W = 3\ \mu$m, unless otherwise specified. Find the labeled currents and voltages.

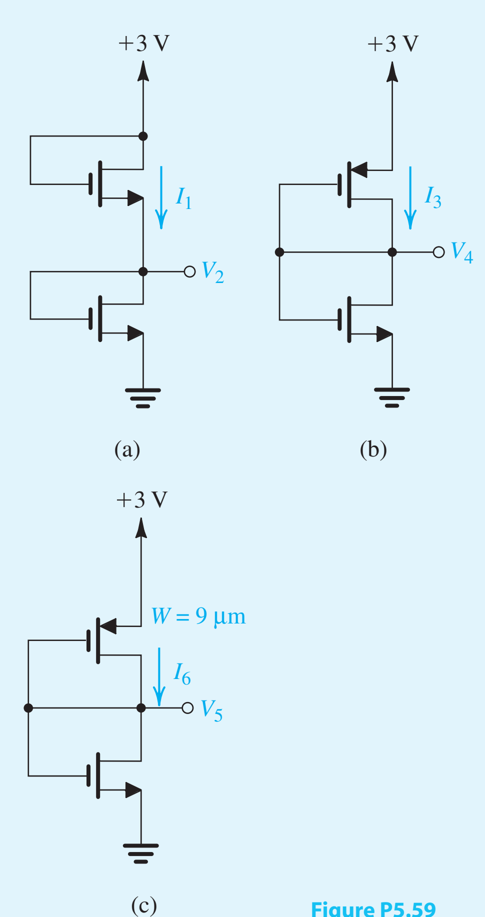


Figure P5.59

SIM ***5.60** For the devices in the circuit of Fig. P5.60, $|V_t| = 1$ V, $\lambda = 0$, $\mu_n C_{ox} = 50\ \mu\text{A/V}^2$, $L = 1\ \mu\text{m}$, and $W = 10\ \mu\text{m}$. Find V_2 and I_2. How do these values change if Q_3 and Q_4 are made to have $W = 100\ \mu\text{m}$?

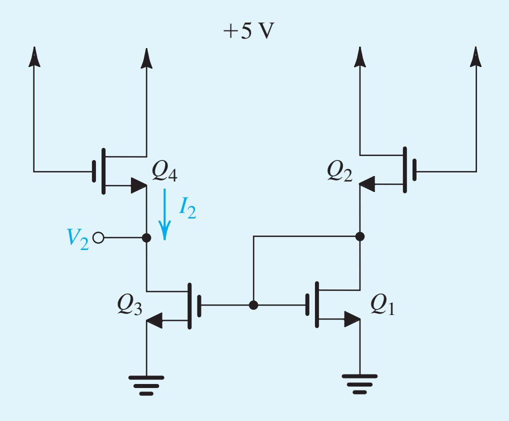


Figure P5.60

5.61 In the circuit of Fig. P5.61, transistors Q_1 and Q_2 have $V_t = 0.7$ V, and the process transconductance parameter $k_n' = 125\ \mu\text{A/V}^2$. Find V_1, V_2, and V_3 for each of the following cases:

(a) $(W/L)_1 = (W/L)_2 = 20$
(b) $(W/L)_1 = 1.5(W/L)_2 = 20$

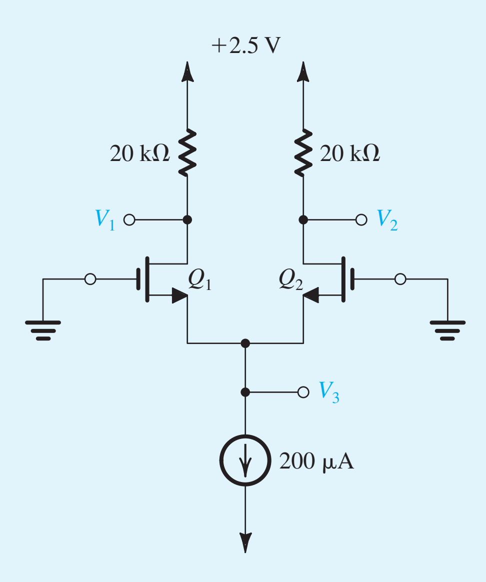


Figure P5.61

Section 5.4: The Body Effect and Other Topics

5.62 In a particular application, an *n*-channel MOSFET operates with V_{SB} in the range 0 V to 4 V. If V_{t0} is nominally 1.0 V, find the range of V_t that results if $\gamma = 0.5\ \text{V}^{1/2}$ and $2\phi_f = 0.6$ V. If the gate oxide thickness is increased by a factor of 4, what does the threshold voltage become?

5.63 A *p*-channel transistor operates in saturation with its source voltage 3 V lower than its substrate. For $\gamma = 0.5\ \text{V}^{1/2}$, $2\phi_f = 0.75$ V, and $V_{t0} = -0.7$ V, find V_t.

***5.64** (a) Using the expression for i_D in saturation and neglecting the channel-length modulation effect (i.e., let $\lambda = 0$), derive an expression for the per unit change in i_D per °C $\left[(\partial i_D/i_D)/\partial T\right]$ in terms of the per unit change in k_n' per °C $\left[(\partial k_n'/k_n')/\partial T\right]$, the temperature coefficient of V_t in V/°C $(\partial V_t/\partial T)$, and V_{GS} and V_t.
(b) If V_t decreases by 2 mV for every °C rise in temperature, find the temperature coefficient of k_n' that results in i_D decreasing by 0.2%/°C when the NMOS transistor with $V_t = 1$ V is operated at $V_{GS} = 5$ V.

5.65 A depletion-type *n*-channel MOSFET with $k_n'W/L = 2\ \text{mA/V}^2$ and $V_t = -3$ V has its source and gate grounded. Find the region of operation and the drain current for $v_D = 0.1$ V, 1 V, 3 V, and 5 V. Neglect the channel-length-modulation effect.

5.66 For a particular depletion-mode NMOS device, $V_t = -2$ V, $k_n'W/L = 200\ \mu\text{A/V}^2$, and $\lambda = 0.02\ \text{V}^{-1}$. When operated at $v_{GS} = 0$, what is the drain current that flows for $v_{DS} = 1$ V, 2 V, 3 V, and 10 V? What does each of these currents become if the device width is doubled with L the same? With L also doubled?

***5.67** Neglecting the channel-length-modulation effect, show that for the depletion-type NMOS transistor of Fig. P5.67, the i–v relationship is given by

$$i = \frac{1}{2}k_n'(W/L)\left(v^2 - 2V_t v\right) \qquad \text{for } v \geq V_t$$
$$i = -\frac{1}{2}k_n'(W/L)V_t^2 \qquad \text{for } v \leq V_t$$

(Recall that V_t is negative.) Sketch the i–v relationship for the case: $V_t = -2$ V and $k_n'(W/L) = 2\ \text{mA/V}^2$.

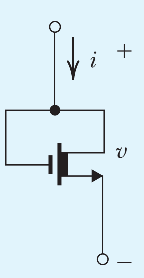


Figure P5.67

CHAPTER 6

Bipolar Junction Transistors (BJTs)

IN THIS CHAPTER YOU WILL LEARN

1. The physical structure of the bipolar transistor and how it works.
2. How the voltage between two terminals of the transistor controls the current that flows through the third terminal, and the equations that describe these current–voltage characteristics.
3. How to analyze and design circuits that contain bipolar transistors, resistors, and dc sources.

Introduction

In this chapter, we study the other major three-terminal device: the bipolar junction transistor (BJT). The presentation of the material in this chapter parallels but does not rely on that for the MOSFET in Chapter 5; thus, if desired, the BJT can be studied before the MOSFET.

Three-terminal devices are far more useful than two-terminal ones, such as the diodes studied in Chapter 4, because they can be used in a multitude of applications, ranging from signal amplification to the design of digital logic and memory circuits. The basic principle involved is the use of the voltage between two terminals to control the current flowing in the third terminal. In this way, a three-terminal device can be used to realize a controlled source, which as we learned in Chapter 1 is the basis for amplifier design. Also, in the extreme, the control signal can be used to cause the current in the third terminal to change from zero to a large value, thus allowing the device to act as a switch. The switch is the basis for the realization of the logic inverter, the basic element of digital circuits.

The invention of the BJT in 1948 at the Bell Telephone Laboratories ushered in the era of solid-state circuits. The result was not just the replacement of vacuum tubes by transistors in radios and television sets but the eruption of an electronics revolution that led to major changes in the way we work, play, and indeed, live. The invention of the transistor also eventually led to the dominance of information technology and the emergence of the knowledge-based economy.

The bipolar transistor enjoyed nearly three decades as the device of choice in the design of both discrete and integrated circuits. Although the MOSFET had been known very early on, it was not until the 1970s and 1980s that it became a serious competitor to the BJT. By 2014, the MOSFET was undoubtedly the most widely used electronic device, and CMOS technology the technology of choice in the design of integrated circuits. Nevertheless, the BJT remains a significant device that excels in certain applications.

The BJT remains popular in discrete-circuit design, where it is used together with other discrete components such as resistors and capacitors to implement circuits that are assembled

on printed-circuit boards (PCBs). Here we note the availability of a very wide selection of BJT types that fit nearly every conceivable application. As well, the BJT is still the preferred device in some very demanding analog and digital integrated-circuit applications. This is especially true in very-high-frequency and high-speed circuits. In particular, a very-high-speed digital logic-circuit family based on bipolar transistors, namely, emitter-coupled logic, is still in use (Chapter 15). Finally, bipolar transistors can be combined with MOSFETs to create innovative circuits that take advantage of the high-input-impedance and low-power operation of MOSFETs and the very-high-frequency operation and high-current-driving capability of bipolar transistors. The resulting technology is known as BiCMOS, and it is finding increasingly larger areas of application (see Chapters 8, 9, 13, and 15).

In this chapter, we shall start with a description of the physical operation of the BJT. Though simple, this physical description provides considerable insight regarding the performance of the transistor as a circuit element. We then quickly move from describing current flow in terms of electrons and holes to a study of the transistor terminal characteristics. Circuit models for transistor operation in different modes will be developed and utilized in the analysis and design of transistor circuits. The main objective of this chapter is to develop in the reader a high degree of familiarity with the BJT. Thus, it lays the foundation for the use of the BJT in amplifier design (Chapter 7).

6.1 Device Structure and Physical Operation

6.1.1 Simplified Structure and Modes of Operation

Figure 6.1 shows a simplified structure for the BJT. A practical transistor structure will be shown later (see also Appendix A, which deals with fabrication technology).

As shown in Fig. 6.1, the BJT consists of three semiconductor regions: the emitter region (*n* type), the base region (*p* type), and the collector region (*n* type). Such a transistor is called an *npn* transistor. Another transistor, a dual of the *npn* as shown in Fig. 6.2, has a *p*-type emitter, an *n*-type base, and a *p*-type collector, and is appropriately called a *pnp* transistor.

A terminal is connected to each of the three semiconductor regions of the transistor, with the terminals labeled **emitter** (E), **base** (B), and **collector** (C).

The transistor consists of two *pn* junctions, the **emitter–base junction** (EBJ) and the **collector–base junction** (CBJ). Depending on the bias condition (forward or reverse) of each of these junctions, different modes of operation of the BJT are obtained, as shown in Table 6.1. The **active mode** is the one used if the transistor is to operate as an amplifier. Switching applications (e.g., logic circuits) utilize both the **cutoff mode** and the **saturation mode**. As the name implies, in the cutoff mode no current flows because both junctions are reverse biased.

As we will see shortly, charge carriers of both polarities—that is, electrons and holes—participate in the current-conduction process in a bipolar transistor, which is the reason for the name *bipolar*.[1]

[1]This should be contrasted with the situation in the MOSFET, where current is conducted by charge carriers of one type only: electrons in *n*-channel devices or holes in *p*-channel devices. In earlier days, some referred to FETs as unipolar devices.

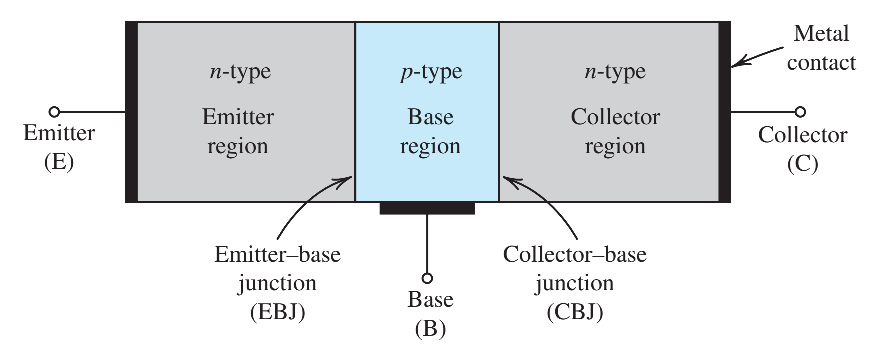


Figure 6.1 A simplified structure of the *npn* transistor.

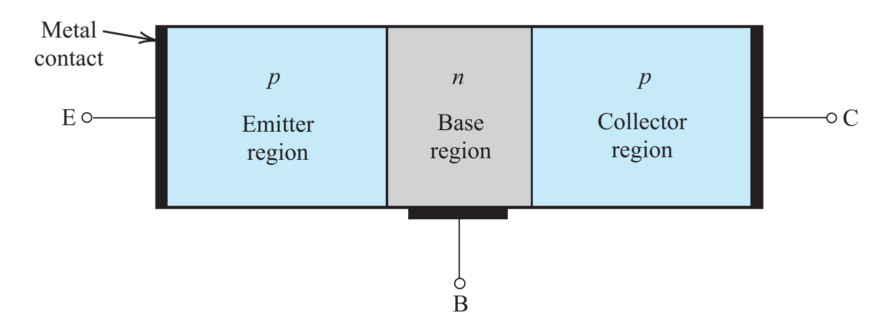


Figure 6.2 A simplified structure of the *pnp* transistor.

Table 6.1 BJT Modes of Operation

Mode	EBJ	CBJ
Cutoff	Reverse	Reverse
Active	Forward	Reverse
Saturation	Forward	Forward

6.1.2 Operation of the *npn* Transistor in the Active Mode

Of the three modes of operation of the BJT, the active mode is the most important. Therefore, we begin our study of the BJT by considering its physical operation in the active mode.[2] This situation is illustrated in Fig. 6.3 for the *npn* transistor. Two external voltage sources (shown as batteries) are used to establish the required bias conditions for active-mode operation. The voltage V_{BE} causes the *p*-type base to be higher in potential than the *n*-type emitter, thus forward biasing the emitter–base junction. The collector–base voltage V_{CB} causes the *n*-type collector to be at a higher potential than the *p*-type base, thus reverse biasing the collector–base junction.

[2]The material in this section assumes that the reader is familiar with the operation of the *pn* junction under forward-bias conditions (Section 3.5).

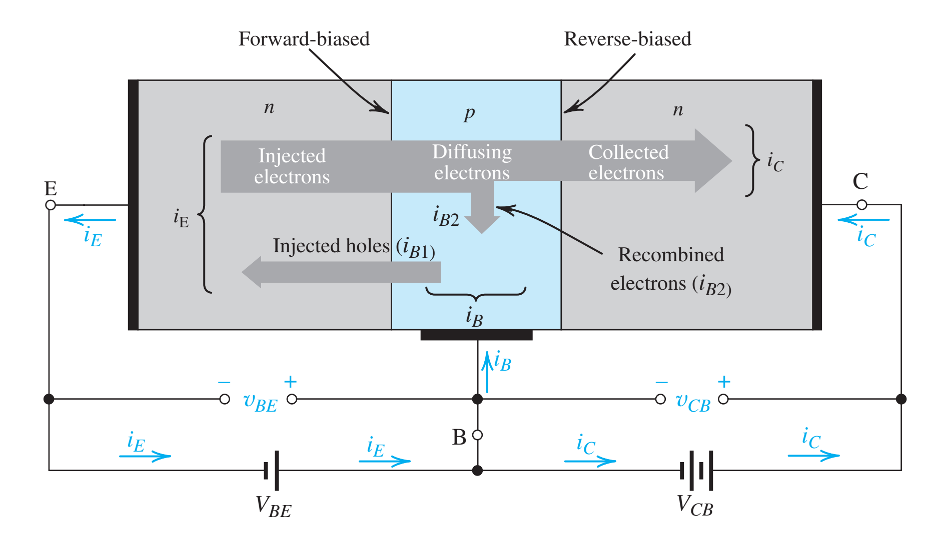


Figure 6.3 Current flow in an *npn* transistor biased to operate in the active mode. (Reverse current components due to drift of thermally generated minority carriers are not shown.)

Current Flow The forward bias on the emitter–base junction will cause current to flow across this junction. Current will consist of two components: electrons injected from the emitter into the base, and holes injected from the base into the emitter. As will become apparent shortly, it is highly desirable to have the first component (electrons from emitter to base) be much larger than the second component (holes from base to emitter). This can be accomplished by fabricating the device with a heavily doped emitter and a lightly doped base; that is, the device is designed to have a high density of electrons in the emitter and a low density of holes in the base.

The current that flows across the emitter–base junction will constitute the emitter current i_E, as indicated in Fig. 6.3. The direction of i_E is "out of" the emitter lead, which, following the usual conventions, is in the direction of the positive-charge flow (hole current) and opposite to the direction of the negative-charge flow (electron current), with the emitter current i_E being equal to the sum of these two components. However, since the electron component is much larger than the hole component, the emitter current will be dominated by the electron component.

From our study in Section 3.5 of the current flow across a forward-biased *pn* junction, we know that the magnitude of both the electron component and the hole component of i_E will be proportional to e^{v_{BE}/V_T}, where v_{BE} is the forward voltage across the base–emitter junction and V_T is the thermal voltage (approximately 25 mV at room temperature).

Let's now focus our attention on the first current component, namely, that carried by electrons injected from the emitter into the base. These electrons will be **minority carriers** in the *p*-type base region. Because their concentration will be highest at the emitter side of the base, the injected electrons will diffuse through the base region toward the collector. In their journey across the base, some of the electrons will combine with holes, which are majority carriers in the base. However, since the base is usually very thin and, as mentioned earlier, lightly doped, the proportion of electrons that are "lost" through this **recombination process** will be quite small. Thus, most of the diffusing electrons will reach the boundary of the collector–base depletion region. Because the collector is more positive than the base (by the

reverse-bias voltage v_{CB}), these successful electrons will be swept across the CBJ depletion region into the collector. They will thus get collected and constitute the collector current i_C.

The Collector Current From the foregoing statements, we see that the collector current is carried by the electrons that reach the collector region. Its direction will be opposite to that of the flow of electrons, and thus into the collector terminal. Its magnitude will be proportional to e^{v_{BE}/V_T}, thus

$$i_C = I_S e^{v_{BE}/V_T} \tag{6.1}$$

where the constant of proportionality I_S, as in the case of the diode, is called the **saturation current** and is a transistor parameter. We will have more to say about I_S shortly.

An important observation to make here is that i_C is independent of the value of v_{CB}. That is, as long as the collector is positive with respect to the base, the electrons that reach the collector side of the base region will be swept into the collector and will register as collector current.

The Base Current Reference to Fig. 6.3 shows that the base current i_B is composed of two components. The first component i_{B1} is due to the holes injected from the base region into the emitter region. This current component is proportional to e^{v_{BE}/V_T}. The second component of base current, i_{B2}, is due to holes that have to be supplied by the external circuit in order to replace the holes lost from the base through the recombination process. Because i_{B2} is proportional to the number of electrons injected into the base, it also will be proportional to e^{v_{BE}/V_T}. Thus the total base current, $i_B = i_{B1} + i_{B2}$, will be proportional to e^{v_{BE}/V_T}, and can be expressed as a fraction of the collector current i_C as follows:

$$i_B = \frac{i_C}{\beta} \tag{6.2}$$

That is,

$$i_B = \left(\frac{I_S}{\beta}\right) e^{v_{BE}/V_T} \tag{6.3}$$

where β is a transistor parameter.

For modern *npn* transistors, β is in the range 50 to 200, but it can be as high as 1000 for special devices. For reasons that will become clear later, the parameter β is called the **common-emitter current gain**.

The above description indicates that the value of β is highly influenced by two factors: the width of the base region, W, and the relative dopings of the base region and the emitter region, N_A/N_D. To obtain a high β (which is highly desirable since β represents a gain parameter) the base should be thin (W small) and lightly doped and the emitter heavily doped (making N_A/N_D small). For modern integrated circuit fabrication technologies, W is in the nanometer range.

The Emitter Current Since the current that enters a transistor must leave it, it can be seen from Fig. 6.3 that the emitter current i_E is equal to the sum of the collector current i_C and the base current i_B; that is,

$$i_E = i_C + i_B \tag{6.4}$$

Use of Eqs. (6.2) and (6.4) gives

$$i_E = \frac{\beta + 1}{\beta} i_C \tag{6.5}$$

That is,

$$i_E = \frac{\beta + 1}{\beta} I_S e^{v_{BE}/V_T} \tag{6.6}$$

Alternatively, we can express Eq. (6.5) in the form

$$i_C = \alpha i_E \tag{6.7}$$

where the constant α is related to β by

$$\alpha = \frac{\beta}{\beta + 1} \tag{6.8}$$

Thus the emitter current in Eq. (6.6) can be written

$$i_E = (I_S/\alpha) e^{v_{BE}/V_T} \tag{6.9}$$

Finally, we can use Eq. (6.8) to express β in terms of α, that is,

$$\beta = \frac{\alpha}{1 - \alpha} \tag{6.10}$$

It can be seen from Eq. (6.8) that α is a constant (for a particular transistor) that is less than but very close to unity. For instance, if $\beta = 100$, then $\alpha \simeq 0.99$. Equation (6.10) reveals an important fact: Small changes in α correspond to very large changes in β. This mathematical observation manifests itself physically, with the result that transistors of the same type may have widely different values of β. For reasons that will become apparent later, α is called the **common-base current gain**.

Minority-Carrier Distribution Our understanding of the physical operation of the BJT can be enhanced by considering the distribution of minority charge carriers in the base and the emitter. Figure 6.4 shows the profiles of the concentration of electrons in the base and holes in the emitter of an *npn* transistor operating in the active mode. Observe that since the doping concentration in the emitter, N_D, is much higher than the doping concentration in the base, N_A, the concentration of electrons injected from emitter to base, $n_p(0)$, is much higher than the concentration of holes injected from the base to the emitter, $p_n(0)$. Both quantities are proportional to e^{v_{BE}/V_T}, thus

$$n_p(0) = n_{p0} e^{v_{BE}/V_T} \tag{6.11}$$

where n_{p0} is the thermal-equilibrium value of the minority-carrier (electron) concentration in the base region.

Next, observe that because the base is very thin, the concentration of excess electrons decays almost linearly (as opposed to the usual exponential decay, as observed for the excess holes in the emitter region). Furthermore, the reverse bias on the collector–base junction causes the concentration of excess electrons at the collector side of the base to be zero. (Recall that electrons that reach that point are swept into the collector.)

The tapered minority-carrier concentration profile (Fig. 6.4) causes the electrons injected into the base to diffuse through the base region toward the collector. This electron diffusion

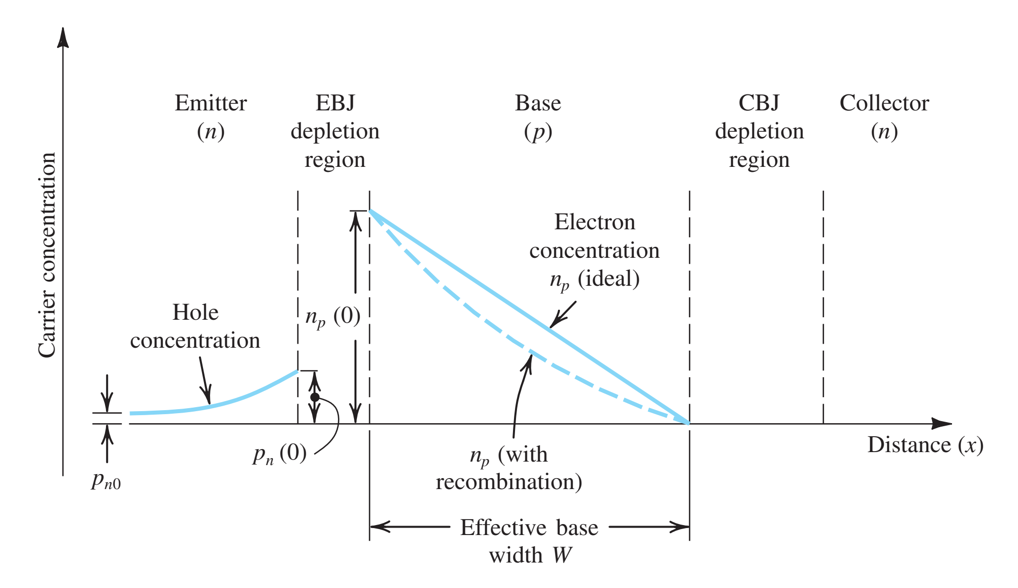


Figure 6.4 Profiles of minority-carrier concentrations in the base and in the emitter of an *npn* transistor operating in the active mode: $v_{BE} > 0$ and $v_{CB} \geq 0$.

current I_n is directly proportional to the slope of the straight-line concentration profile,

$$\begin{aligned} I_n &= A_E q D_n \frac{dn_p(x)}{dx} \\ &= A_E q D_n \left(-\frac{n_p(0)}{W}\right) \end{aligned} \tag{6.12}$$

where A_E is the cross-sectional area of the base–emitter junction (in the direction perpendicular to the page), q is the magnitude of the electron charge, D_n is the electron diffusivity in the base, and W is the effective width of the base. Observe that the negative slope of the minority-carrier concentration results in a negative current I_n across the base; that is, I_n flows from right to left (in the negative direction of x), which corresponds to the usual convention, namely, opposite to the direction of electron flow.

The recombination in the base region, though slight, causes the excess minority-carrier concentration profile to deviate from a straight line and take the slightly concave shape indicated by the broken line in Fig. 6.4. The slope of the concentration profile at the EBJ is slightly higher than that at the CBJ, with the difference accounting for the small number of electrons lost in the base region through recombination.

Finally, we have the collector current $i_C = I_n$, which will yield a negative value for i_C, indicating that i_C flows in the negative direction of the x axis (i.e., from right to left). Since we will take this to be the positive direction of i_C, we can drop the negative sign in Eq. (6.12). Doing this and substituting for $n_p(0)$ from Eq. (6.11), we can thus express the collector current i_C as

$$i_C = I_S e^{v_{BE}/V_T}$$

where the saturation current I_S is given by

$$I_S = A_E q D_n n_{p0}/W$$

Substituting $n_{p0} = n_i^2/N_A$, where n_i is the intrinsic carrier density and N_A is the doping concentration in the base, we can express I_S as

$$I_S = \frac{A_E q D_n n_i^2}{N_A W} \tag{6.13}$$

The saturation current I_S is inversely proportional to the base width W and is directly proportional to the area of the EBJ. Typically I_S is in the range of 10^{-12} A to 10^{-18} A (depending on the size of the device). Because I_S is proportional to n_i^2, it is a strong function of temperature, approximately doubling for every 5°C rise in temperature. (For the dependence of n_i^2 on temperature, refer to Eq. 3.2.)

Since I_S is directly proportional to the junction area (i.e., the device size), it will also be referred to as the **scale current**. Two transistors that are identical except that one has an EBJ area, say, twice that of the other will have saturation currents with that same ratio (i.e., 2). Thus for the same value of v_{BE} the larger device will have a collector current twice that in the smaller device. This concept is frequently employed in integrated-circuit design.

Recapitulation and Equivalent-Circuit Models We have presented a first-order model for the operation of the *npn* transistor in the active mode. Basically, the forward-bias voltage v_{BE} causes an exponentially related current i_C to flow in the collector terminal. The collector current i_C is independent of the value of the collector voltage as long as the collector–base junction remains reverse biased; that is, $v_{CB} \geq 0$. Thus in the active mode the collector terminal behaves as an ideal constant-current source where the value of the current is determined by v_{BE}. The base current i_B is a factor $1/\beta$ of the collector current, and the emitter current is equal to the sum of the collector and base currents. Since i_B is much smaller than i_C (i.e., $\beta \gg 1$), $i_E \simeq i_C$. More precisely, the collector current is a fraction α of the emitter current, with α smaller than, but close to, unity.

This first-order model of transistor operation in the active mode can be represented by the equivalent circuit shown in Fig. 6.5(a). Here, diode D_E has a scale current I_{SE} equal to (I_S/α) and thus provides a current i_E related to v_{BE} according to Eq. (6.9). The current of the controlled source, which is equal to the collector current, is controlled by v_{BE} according to the exponential relationship indicated, a restatement of Eq. (6.1). This model is in essence a nonlinear voltage-controlled current source. It can be converted to the current-controlled current-source model shown in Fig. 6.5(b) by expressing the current of the controlled source as αi_E. Note that this model is also nonlinear because of the exponential relationship of the current i_E through diode D_E and the voltage v_{BE}. From this model we observe that if the transistor is used as a two-port network with the input port between E and B and the output port between C and B (i.e., with B as a common terminal), then the current gain observed is equal to α. Thus α is called the common-base current gain.

Two other equivalent-circuit models, shown in Fig. 6.5(c) and (d), may be used to represent the operation of the BJT. The model of Fig. 6.5(c) is essentially a voltage-controlled current source. However, here diode D_B conducts the base current and thus its current scale factor is I_S/β, resulting in the i_B–v_{BE} relationship given in Eq. (6.3). By simply expressing the collector current as βi_B we obtain the current-controlled current-source model shown in Fig. 6.5(d). From this latter model we observe that if the transistor is used as a two-port network with the input port between B and E and the output port between C and E (i.e., with E as the common terminal), then the current gain observed is equal to β. Thus β is called the common-emitter current gain.

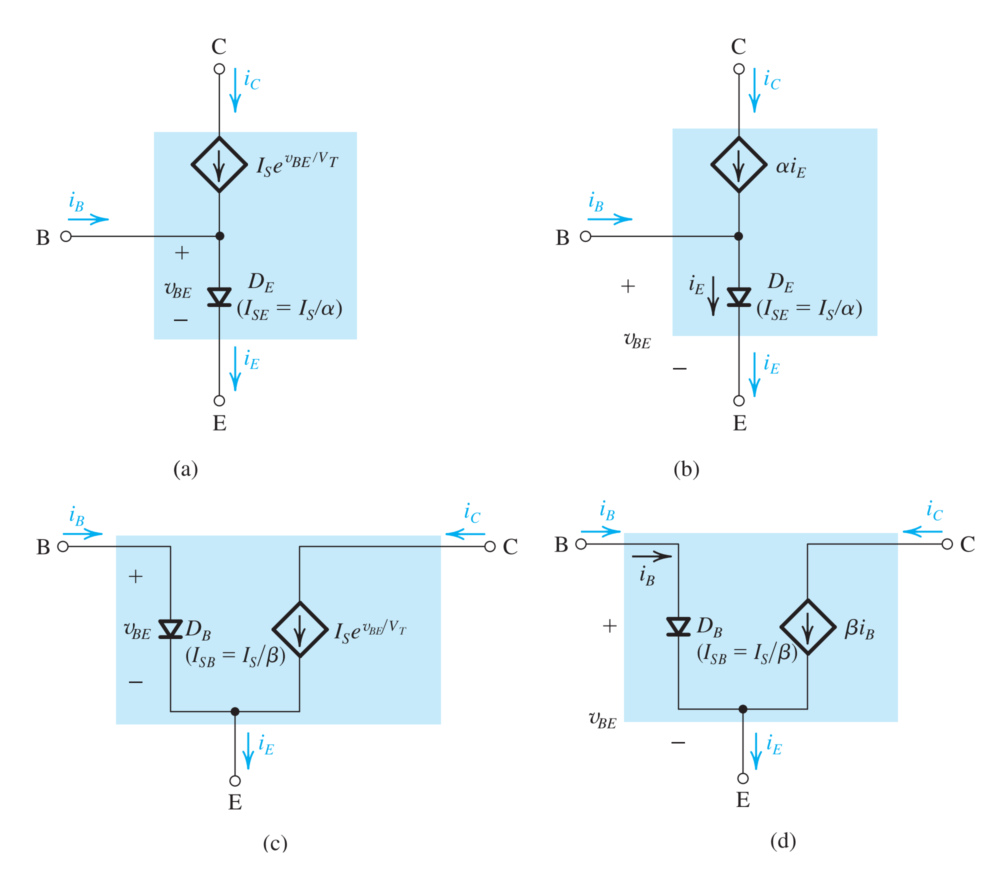


Figure 6.5 Large-signal equivalent-circuit models of the *npn* BJT operating in the forward active mode.

Finally, we note that the models in Fig. 6.5 apply for any positive value of v_{BE}. That is, unlike the models we will be discussing in Chapter 7, here there is no limitation on the size of v_{BE}, and thus these models are referred to as **large-signal models**.

Example 6.1

An *npn* transistor having $I_S = 10^{-15}$ A and $\beta = 100$ is connected as follows: The emitter is grounded, the base is fed with a constant-current source supplying a dc current of 10 μA, and the collector is connected to a 5-V dc supply via a resistance R_C of 3 kΩ. Assuming that the transistor is operating in the active mode, find V_{BE} and V_{CE}. Use these values to verify active-mode operation. Replace the current source with a resistance connected from the base to the 5-V dc supply. What resistance value is needed to result in the same operating conditions?

Example 6.1 *continued*

Solution

If the transistor is operating in the active mode, it can be represented by one of the four possible equivalent-circuit models shown in Fig. 6.5. Because the emitter is grounded, either the model in Fig. 6.5(c) or that in Fig. 6.5(d) would be suitable. Since we know the base current I_B, the model of Fig. 6.5(d) is the most suitable.

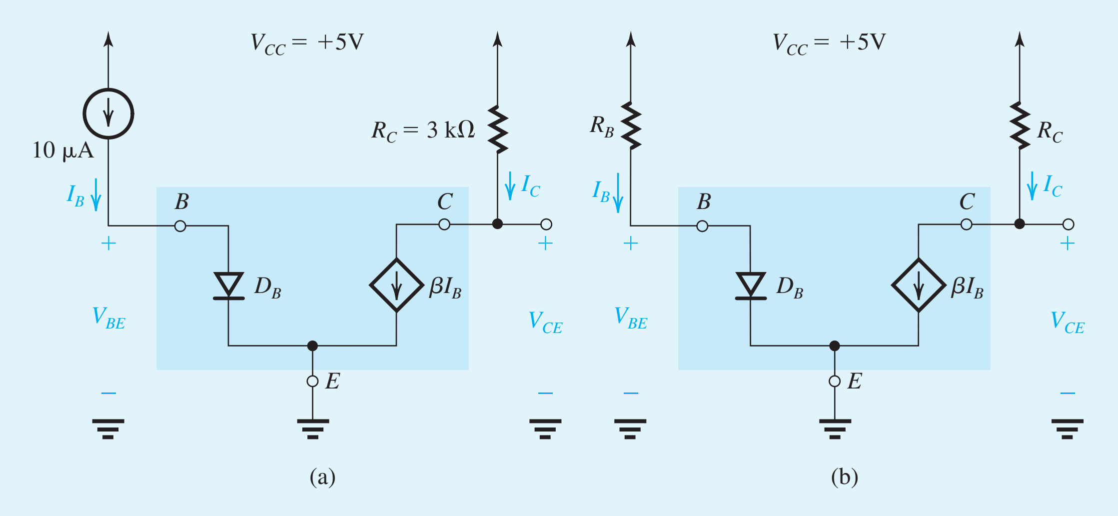


Figure 6.6 Circuits for Example 6.1.

Figure 6.6(a) shows the circuit as described with the transistor represented by the model of Fig. 6.5(d). We can determine V_{BE} from the exponential characteristic of D_B as follows:

$$V_{BE} = V_T \ln \frac{I_B}{I_S/\beta}$$
$$= 25 \ln \left(\frac{10 \times 10^{-6}}{10^{-17}} \right)$$
$$= 690 \text{ mV} = 0.69 \text{ V}$$

Next we determine the value of V_{CE} from

$$V_{CE} = V_{CC} - R_C I_C$$

where

$$I_C = \beta I_B = 100 \times 10 \times 10^{-6} = 10^{-3} \text{ A} = 1 \text{ mA}$$

Thus,

$$V_{CE} = 5 - 3 \times 1 = +2 \text{ V}$$

Since V_C at +2 V is higher than V_B at 0.69 V, the transistor is indeed operating in the active mode.

Now, replacing the 10-μA current source with a resistance R_B connected from the base to the 5-V dc supply V_{CC}, as in Fig. 6.6(b), the value of R_B must be

$$R_B = \frac{V_{CC} - V_{BE}}{I_B}$$

$$= \frac{5 - 0.69}{10\ \mu\text{A}} = 431\ \text{k}\Omega$$

EXERCISES

6.1 Consider an *npn* transistor with $v_{BE} = 0.7$ V at $i_C = 1$ mA. Find v_{BE} at $i_C = 0.1$ mA and 10 mA.
Ans. 0.64 V; 0.76 V

6.2 Transistors of a certain type are specified to have β values in the range of 50 to 150. Find the range of their α values.
Ans. 0.980 to 0.993

6.3 Measurement of an *npn* BJT in a particular circuit shows the base current to be 14.46 μA, the emitter current to be 1.460 mA, and the base–emitter voltage to be 0.7 V. For these conditions, calculate α, β, and I_S.
Ans. 0.99; 100; 10^{-15} A

6.4 Calculate β for two transistors for which $\alpha = 0.99$ and 0.98. For collector currents of 10 mA, find the base current of each transistor.
Ans. 99; 49; 0.1 mA; 0.2 mA

6.5 A transistor for which $I_S = 10^{-16}$ A and $\beta = 100$ is conducting a collector current of 1 mA. Find v_{BE}. Also, find I_{SE} and I_{SB} for this transistor.
Ans. 747.5 mV; 1.01×10^{-16} A; 10^{-18} A

6.6 For the circuit in Fig. 6.6(a) analyzed in Example 6.1, find the maximum value of R_C that will still result in active-mode operation.
Ans. 4.31 kΩ

6.1.3 Structure of Actual Transistors

Figure 6.7 shows a more realistic (but still simplified) cross section of an *npn* BJT. Note that the collector virtually surrounds the emitter region, thus making it difficult for the electrons injected into the thin base to escape being collected. In this way, the resulting α is close to

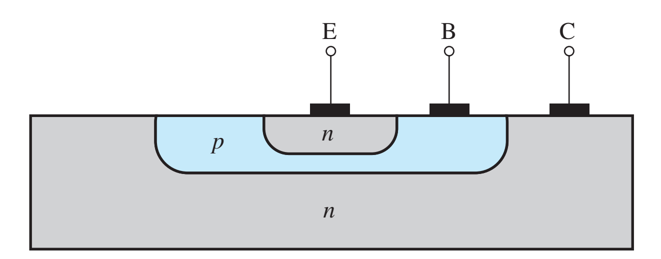


Figure 6.7 Cross section of an *npn* BJT.

unity and β is large. Also, observe that the device is *not* symmetrical, and thus the emitter and collector cannot be interchanged.[3] For more detail on the physical structure of actual devices, the reader is referred to Appendix A.

The structure in Fig. 6.7 indicates also that the CBJ has a much larger area than the EBJ. Thus the CB diode D_C has a saturation current I_{SC} that is much larger than the saturation current of the EB diode D_E. Typically, I_{SC} is 10 to 100 times larger than I_{SE} (recall that $I_{SE} = I_S/\alpha \simeq I_S$).

EXERCISE

6.7 A particular transistor has $I_S = 10^{-15}$ A and $\alpha \simeq 1$. If the CBJ area is 100 times the area of the EBJ, find the collector scale current I_{SC}.
Ans. 10^{-13} A

6.1.4 Operation in the Saturation Mode[4]

As mentioned above, for the BJT to operate in the active mode, the CBJ must be reverse biased. Thus far, we have stated this condition for the *npn* transistor as $v_{CB} \geq 0$. However, we know that a *pn* junction does not effectively become forward biased until the forward voltage across it exceeds approximately 0.4 V. It follows that one can maintain active-mode operation of an *npn* transistor for negative v_{CB} down to approximately −0.4 V. This is illustrated in Fig. 6.8, which is a sketch of i_C versus v_{CB} for an *npn* transistor operated with a constant emitter current I_E. As expected, i_C is independent of v_{CB} in the active mode, a situation that extends

[3]If the emitter and collector are reversed—that is, the CBJ is forward biased and the EBJ is reverse biased—the device operates in a mode called the "reverse-active mode." The resulting values of α and β, denoted α_R and β_R (with R denoting reverse), are much lower than the values of α and β, respectively, obtained in the "forward"-active mode discussed above. Hence, the reverse-active mode has no practical application. The MOSFET, on the other hand, being a perfectly symmetrical device, can operate equally well with its drain and source terminals interchanged.

[4]Saturation means something completely different in a BJT and in a MOSFET. The saturation mode of operation of the BJT is analogous to the triode region of operation of the MOSFET. On the other hand, the saturation region of operation of the MOSFET corresponds to the active mode of BJT operation.

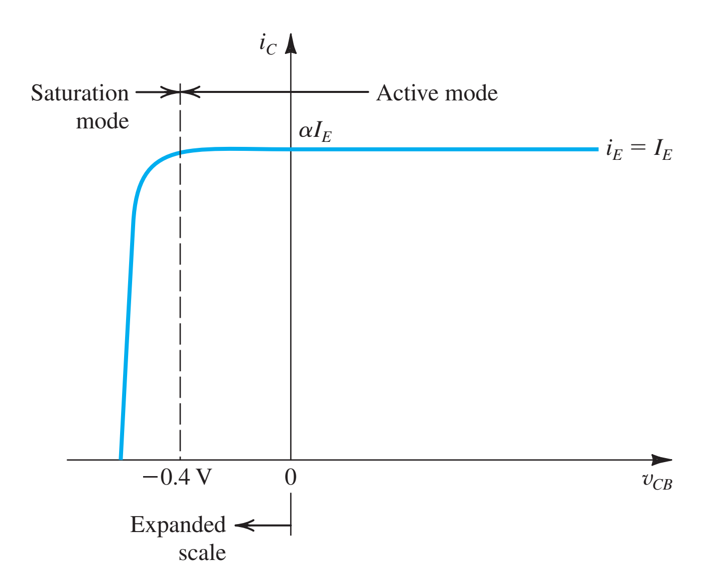


Figure 6.8 The i_C–v_{CB} characteristic of an *npn* transistor fed with a constant emitter current I_E. The transistor enters the saturation mode of operation for $v_{CB} < -0.4$ V, and the collector current diminishes.

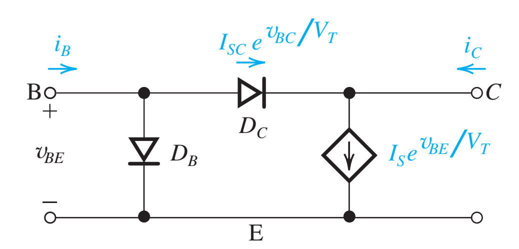


Figure 6.9 Modeling the operation of an *npn* transistor in saturation by augmenting the model of Fig. 6.5(c) with a forward-conducting diode D_C. Note that the current through D_C increases i_B and reduces i_C.

for v_{CB} going negative to approximately −0.4 V. Below this value of v_{CB}, the CBJ begins to conduct sufficiently that the transistor leaves the active mode and enters the saturation mode of operation, where i_C decreases.

To see why i_C decreases in saturation, we can construct a model for the saturated *npn* transistor as follows. We augment the model of Fig. 6.5(c) with the forward-conducting CBJ diode D_C, as shown in Fig. 6.9. Observe that the current i_{BC} will subtract from the controlled-source current, resulting in the reduced collector current i_C given by

$$i_C = I_S e^{v_{BE}/V_T} - I_{SC} e^{v_{BC}/V_T} \tag{6.14}$$

where I_{SC} is the saturation current for D_C and is related to I_S by the ratio of the areas of the CBJ and the EBJ. The second term in Eq. (6.14) will play an increasing role as v_{BC} exceeds 0.4 V or so, causing i_C to decrease and eventually reach zero.

Figure 6.9 also indicates that in saturation the base current will increase to the value

$$i_B = (I_S/\beta) e^{v_{BE}/V_T} + I_{SC} e^{v_{BC}/V_T} \tag{6.15}$$

Equations (6.14) and (6.15) can be combined to obtain the ratio i_C/i_B for a saturated transistor. We observe that this ratio will be *lower* than the value of β. Furthermore, the ratio will decrease as v_{BC} is increased and the transistor is driven deeper into saturation. Because i_C/i_B

of a saturated transistor can be set to any desired value lower than β by adjusting v_{BC}, this ratio is known as **forced** β and denoted β_{forced},

$$\beta_{\text{forced}} = \left.\frac{i_C}{i_B}\right|_{\text{saturation}} \leq \beta \tag{6.16}$$

As will be shown later, in analyzing a circuit we can determine whether the BJT is in the saturation mode by either of the following two tests:

1. Is the CBJ forward biased by more than 0.4 V?
2. Is the ratio i_C/i_B lower than β?

The collector-to-emitter voltage v_{CE} of a saturated transistor can be found from Fig. 6.9 as the difference between the forward-bias voltages of the EBJ and the CBJ,

$$V_{CE\text{sat}} = V_{BE} - V_{BC} \tag{6.17}$$

Recalling that the CBJ has a much larger area than the EBJ, V_{BC} will be smaller than V_{BE} by 0.1 to 0.3 V. Thus,



$$V_{CE\text{sat}} \simeq 0.1 \text{ to } 0.3\text{ V}$$

Typically we will assume that a transistor at the edge of saturation has $V_{CE\text{sat}} = 0.3$ V, while a transistor deep in saturation has $V_{CE\text{sat}} = 0.2$ V.

EXERCISES

6.8 Use Eq. (6.14) to show that i_C reaches zero at

$$V_{CE} = V_T \ln\left(I_{SC}/I_S\right)$$

Calculate V_{CE} for a transistor whose CBJ has 100 times the area of the EBJ.
Ans. 115 mV

6.9 Use Eqs. (6.14), (6.15), and (6.16) to show that a BJT operating in saturation with $V_{CE} = V_{CE\text{sat}}$ has a forced β given by

$$\beta_{\text{forced}} = \beta \frac{e^{V_{CE\text{sat}}/V_T} - I_{SC}/I_S}{e^{V_{CE\text{sat}}/V_T} + \beta I_{SC}/I_S}$$

Find β_{forced} for $\beta = 100$, $I_{SC}/I_S = 100$, and $V_{CE\text{sat}} = 0.2$ V.
Ans. 22.2

6.1.5 The *pnp* Transistor

The *pnp* transistor operates in a manner similar to that of the *npn* device described above. Figure 6.10 shows a *pnp* transistor biased to operate in the active mode. Here the voltage V_{EB} causes the *p*-type emitter to be higher in potential than the *n*-type base, thus forward biasing the emitter–base junction. The collector–base junction is reverse biased by the voltage V_{BC}, which keeps the *p*-type collector lower in potential than the *n*-type base.

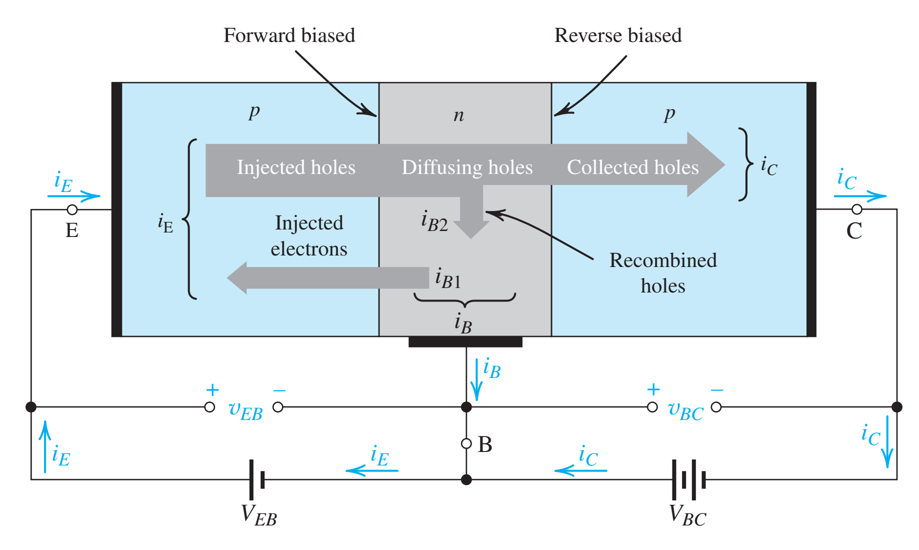


Figure 6.10 Current flow in a *pnp* transistor biased to operate in the active mode.

Unlike the *npn* transistor, current in the *pnp* device is mainly conducted by holes injected from the emitter into the base as a result of the forward-bias voltage V_{EB}. Since the component of emitter current contributed by electrons injected from base to emitter is kept small by using a lightly doped base, most of the emitter current will be due to holes. The electrons injected from base to emitter give rise to the first component of base current, i_{B1}. Also, a number of the holes injected into the base will recombine with the majority carriers in the base (electrons) and will thus be lost. The disappearing base electrons will have to be replaced from the external circuit, giving rise to the second component of base current, i_{B2}. The holes that succeed in reaching the boundary of the depletion region of the collector–base junction will be attracted by the negative voltage on the collector. Thus these holes will be swept across the depletion region into the collector and appear as collector current.

It can easily be seen from the above description that the current–voltage relationship of the *pnp* transistor will be identical to that of the *npn* transistor except that v_{BE} has to be replaced by v_{EB}. Also, the large-signal, active-mode operation of the *pnp* transistor can be modeled by any of four equivalent circuits similar to those for the *npn* transistor in Fig. 6.5. Two of these four circuits are shown in Fig. 6.11. Finally, we note that the *pnp* transistor can operate in the saturation mode in a manner analogous to that described for the *npn* device.

EXERCISES

6.10 Consider the model in Fig. 6.11(a) applied in the case of a *pnp* transistor whose base is grounded, the emitter is fed by a constant-current source that supplies a 2-mA current into the emitter terminal, and the collector is connected to a –10-V dc supply. Find the emitter voltage, the base current, and the collector current if for this transistor $\beta = 50$ and $I_S = 10^{-14}$ A.
Ans. 0.650 V; 39.2 μA; 1.96 mA

6.11 For a *pnp* transistor having $I_S = 10^{-11}$ A and $\beta = 100$, calculate v_{EB} for $i_C = 1.5$ A.
Ans. 0.643 V

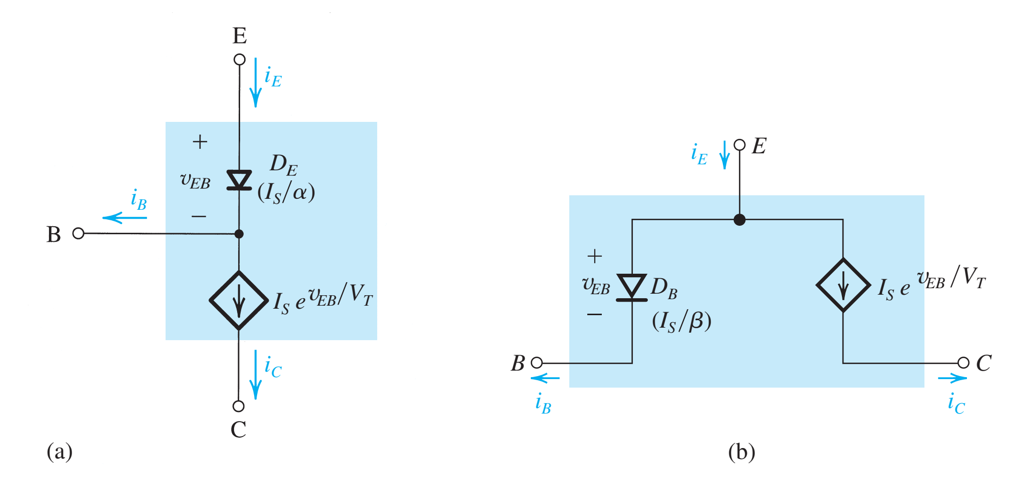


Figure 6.11 Two large-signal models for the *pnp* transistor operating in the active mode.

THE INVENTION OF THE BJT:

The first working transistor was demonstrated at the Bell Labs in late 1947 by John Bardeen and Walter Brattain, who were part of a team led by William Shockley. Made of germanium, the device became known as a point-contact transistor and operated on the field-effect principle. Within a few weeks, however, Shockley wrote a complete description of the bipolar junction transistor (BJT) and filed for a U.S. patent with the title "Circuit Element Utilizing Semiconductor Material."

BJTs dominated the electronics world from the early 1950s to the mid-1970s, when MOSFETs took over the leading position. In 1956, Shockley, Bardeen, and Brattain shared the Nobel Prize in Physics for the discovery of the transistor effect.

6.2 Current–Voltage Characteristics

6.2.1 Circuit Symbols and Conventions

The physical structure used thus far to explain transistor operation is rather cumbersome to employ in drawing the schematic of a multitransistor circuit. Fortunately, a very descriptive and convenient circuit symbol exists for the BJT. Figure 6.12(a) shows the symbol for the *npn* transistor; the *pnp* symbol is given in Fig. 6.12(b). In both symbols the emitter is distinguished by an arrowhead. This distinction is important because, as we have seen in the last section, practical BJTs are not symmetric devices.

The polarity of the device—*npn* or *pnp*—is indicated by the direction of the arrowhead on the emitter. This arrowhead points in the direction of normal current flow in the emitter, which is also the forward direction of the base–emitter junction. Since we have adopted a drawing convention by which currents flow from top to bottom, we will always draw *pnp* transistors in the manner shown in Fig. 6.12(b) (i.e., with their emitters on top).

Figure 6.13 shows *npn* and *pnp* transistors connected to dc sources so as to operate in the active mode. Figure 6.13 also indicates the reference and actual directions of current flow throughout the transistor. Our convention will be to take the reference direction to coincide

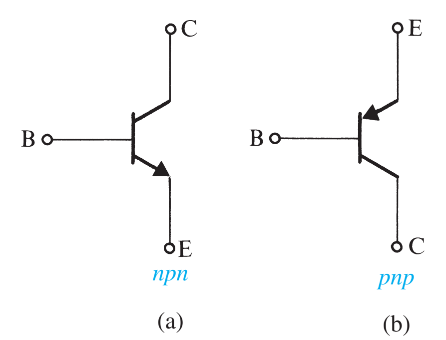


Figure 6.12 Circuit symbols for BJTs.

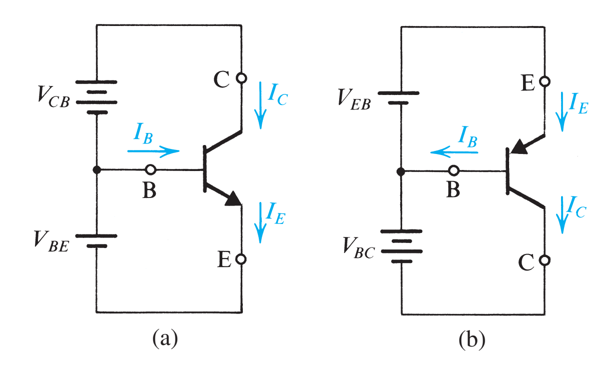


Figure 6.13 Voltage polarities and current flow in transistors operating in the active mode.

with the normal direction of current flow. Hence, normally, we should not encounter a negative value for i_E, i_B, or i_C.

The convenience of the circuit-drawing convention that we have adopted should be obvious from Fig. 6.13. Note that currents flow from top to bottom and that voltages are higher at the top and lower at the bottom. The arrowhead on the emitter also implies the polarity of the emitter–base voltage that should be applied in order to forward bias the emitter–base junction. Just a glance at the circuit symbol of the *pnp* transistor, for example, indicates that we should make the emitter higher in voltage than the base (by v_{EB}) in order to cause current to flow into the emitter (downward). Note that the symbol v_{EB} means the voltage by which the emitter (E) is higher than the base (B). Thus for a *pnp* transistor operating in the active mode v_{EB} is positive, while in an *npn* transistor v_{BE} is positive.

From the discussion of Section 6.1 it follows that an *npn* transistor whose EBJ is forward biased (usually, $V_{BE} \simeq 0.7$ V) will operate in the active mode *as long as the collector voltage does not fall below that of the base by more than approximately 0.4 V*. Otherwise, the transistor leaves the active mode and enters the saturation region of operation.[5]

In a parallel manner, the *pnp* transistor will operate in the active mode *if the EBJ is forward biased (usually, $V_{EB} \simeq 0.7$ V) and the collector voltage is not allowed to rise above that of the base by more than 0.4 V or so*. Otherwise, the CBJ becomes forward biased, and the *pnp* transistor enters the saturation region of operation.

[5]It is interesting to contrast the active-mode operation of the BJT with the corresponding mode of operation of the MOSFET: The BJT needs a minimum v_{CE} of about 0.3 V, and the MOSFET needs a minimum v_{DS} equal to V_{OV}, which for modern technologies is in the range of 0.2 V to 0.3 V. Thus we see a great deal of similarity! Also note that reverse biasing the CBJ of the BJT corresponds to pinching off the channel of the MOSFET. This condition results in the collector current (drain current in the MOSFET) being independent of the collector voltage (the drain voltage in the MOSFET).

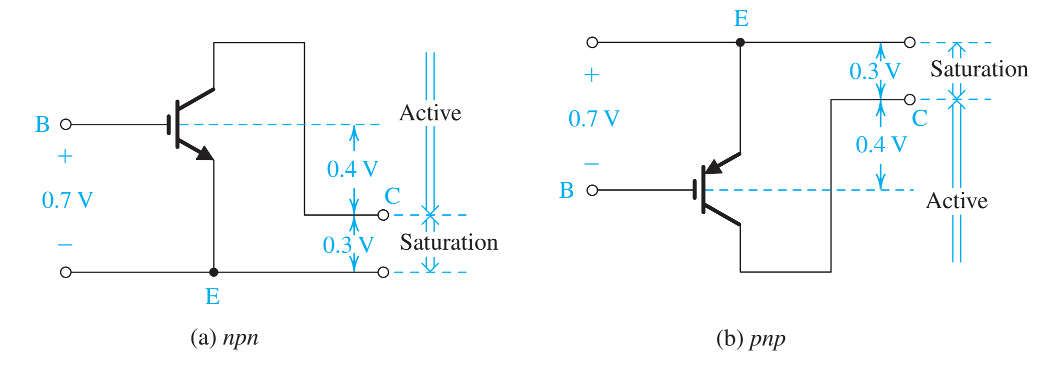


Figure 6.14 Graphical representation of the conditions for operating the BJT in the active mode and in the saturation mode.

Table 6.2 Summary of the BJT Current–Voltage Relationships in the Active Mode

$$i_C = I_S e^{v_{BE}/V_T}$$

$$i_B = \frac{i_C}{\beta} = \left(\frac{I_S}{\beta}\right) e^{v_{BE}/V_T}$$

$$i_E = \frac{i_C}{\alpha} = \left(\frac{I_S}{\alpha}\right) e^{v_{BE}/V_T}$$

Note: For the *pnp* transistor, replace v_{BE} with v_{EB}.

$i_C = \alpha i_E$	$i_B = (1-\alpha)i_E = \dfrac{i_E}{\beta+1}$
$i_C = \beta i_B$	$i_E = (\beta+1)i_B$
$\beta = \dfrac{\alpha}{1-\alpha}$	$\alpha = \dfrac{\beta}{\beta+1}$

$$V_T = \text{thermal voltage} = \frac{kT}{q} \simeq 25 \text{ mV at room temperature}$$

For greater emphasis, we show in Fig. 6.14 a graphical construction that illustrates the conditions for operating the BJT in the active mode and in the saturation mode. Also, for easy reference, we present in Table 6.2 a summary of the BJT current–voltage relationships in the active mode of operation.

The Collector–Base Reverse Current (I_{CBO}) In our discussion of current flow in transistors we ignored the small reverse currents carried by thermally generated minority carriers. Although such currents can be safely neglected in modern transistors, the reverse current across the collector–base junction deserves some mention. This current, denoted I_{CBO}, is the reverse current flowing from collector to base with the emitter open-circuited (hence the subscript O). This current is usually in the nanoampere range, a value that is many times higher than its theoretically predicted value. As with the diode reverse current, I_{CBO} contains a substantial leakage component, and its value is dependent on v_{CB}. I_{CBO} depends strongly on temperature, approximately doubling for every 10°C rise.[6]

[6] The temperature coefficient of I_{CBO} is different from that of I_S because I_{CBO} contains a substantial leakage component.

Example 6.2

The transistor in the circuit of Fig. 6.15(a) has $\beta = 100$ and exhibits a v_{BE} of 0.7 V at $i_C = 1$ mA. Design the circuit so that a current of 2 mA flows through the collector and a voltage of +5 V appears at the collector.

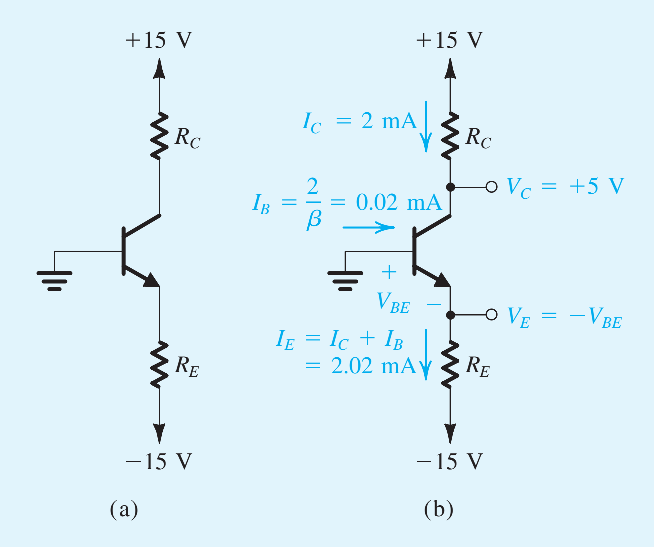


Figure 6.15 Circuit for Example 6.2.

Solution

Refer to Fig. 6.15(b). We note at the outset that since we are required to design for $V_C = +5$ V, the CBJ will be reverse biased and the BJT will be operating in the active mode. To obtain a voltage $V_C = +5$ V, the voltage drop across R_C must be 15 – 5 = 10 V. Now, since $I_C = 2$ mA, the value of R_C should be selected according to

$$R_C = \frac{10\text{ V}}{2\text{ mA}} = 5\text{ k}\Omega$$

Since $v_{BE} = 0.7$ V at $i_C = 1$ mA, the value of v_{BE} at $i_C = 2$ mA is

$$V_{BE} = 0.7 + V_T \ln\left(\frac{2}{1}\right) = 0.717\text{ V}$$

Since the base is at 0 V, the emitter voltage should be

$$V_E = -0.717\text{ V}$$

For $\beta = 100$, $\alpha = 100/101 = 0.99$. Thus the emitter current should be

$$I_E = \frac{I_C}{\alpha} = \frac{2}{0.99} = 2.02\text{ mA}$$

Example 6.2 *continued*

Now the value required for R_E can be determined from

$$R_E = \frac{V_E-(-15)}{I_E}$$
$$= \frac{-0.717+15}{2.02} = 7.07\ \text{k}\Omega$$

This completes the design. We should note, however, that the calculations above were made with a degree of precision that is usually neither necessary nor justified in practice in view, for instance, of the expected tolerances of component values. Nevertheless, we chose to do the design precisely in order to illustrate the various steps involved.

EXERCISES

D6.12 Repeat Example 6.2 for a transistor fabricated in a modern integrated-circuit process. Such a process yields devices that exhibit larger v_{BE} at the same i_C because they have much smaller junction areas. The dc power supplies utilized in modern IC technologies fall in the range of 1 V to 3 V. Design a circuit similar to that shown in Fig. 6.15 except that now the power supplies are ±1.5 V and the BJT has $\beta = 100$ and exhibits v_{BE} of 0.8 V at $i_C = 1$ mA. Design the circuit so that a current of 2 mA flows through the collector and a voltage of +0.5 V appears at the collector.
Ans. $R_C = 500\ \Omega$; $R_E = 338\ \Omega$

6.13 In the circuit shown in Fig. E6.13, the voltage at the emitter was measured and found to be –0.7 V. If $\beta = 50$, find I_E, I_B, I_C, and V_C.

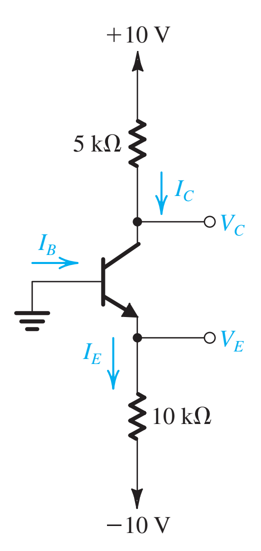


Figure E6.13

Ans. 0.93 mA; 18.2 μA; 0.91 mA; +5.45 V

6.14 In the circuit shown in Fig. E6.14, measurement indicates V_B to be +1.0 V and V_E to be +1.7 V. What are α and β for this transistor? What voltage V_C do you expect at the collector?

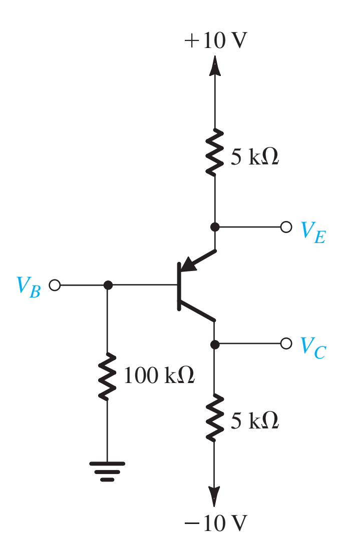


Figure E6.14

Ans. 0.994; 165; –1.75 V

6.2.2 Graphical Representation of Transistor Characteristics

It is sometimes useful to describe the transistor i–v characteristics graphically. Figure 6.16 shows the i_C–v_{BE} characteristic, which is the exponential relationship

$$i_C = I_S e^{v_{BE}/V_T}$$

which is identical to the diode i–v relationship. The $i_E - v_{BE}$ and i_B–v_{BE} characteristics are also exponential but with different scale currents: I_S/α for i_E, and I_S/β for i_B. Since the constant of the exponential characteristic, $1/V_T$, is quite high ($\simeq 40$), the curve rises very sharply. For v_{BE} smaller than about 0.5 V, the current is negligibly small.[7] Also, over most of the normal current range v_{BE} lies in the range of 0.6 V to 0.8 V. In performing rapid first-order dc calculations, we normally will assume that $V_{BE} \simeq 0.7$ V, which is similar to the approach used in the analysis of diode circuits (Chapter 4). For a *pnp* transistor, the $i_C - v_{EB}$ characteristic will look identical to that of Fig. 6.16 with v_{BE} replaced with v_{EB}.

[7]The $i_C - v_{BE}$ characteristic is the BJT's counterpart of the i_D–v_{GS} characteristic of the MOSFET. They share an important attribute: In both cases the voltage has to exceed a "threshold" for the device to conduct appreciably. In the case of the MOSFET, there is a formal threshold voltage, V_t, which lies typically in the range of 0.4 V to 0.8 V. For the BJT, there is an "apparent threshold" of approximately 0.5 V. The i_D–v_{GS} characteristic of the MOSFET is parabolic, and thus is less steep than the i_C–v_{BE} characteristic of the BJT. As will be seen in Chapter 7, this difference has a direct and significant implication for the value of transconductance g_m realized with each device.

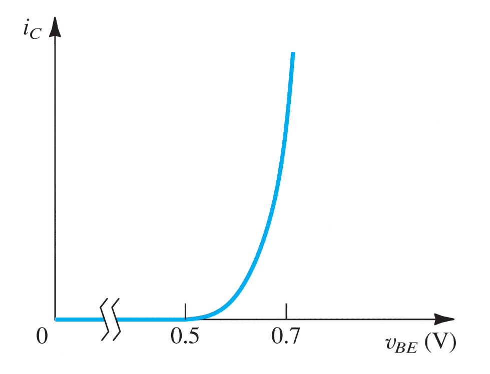


Figure 6.16 The i_C–v_{BE} characteristic for an *npn* transistor.

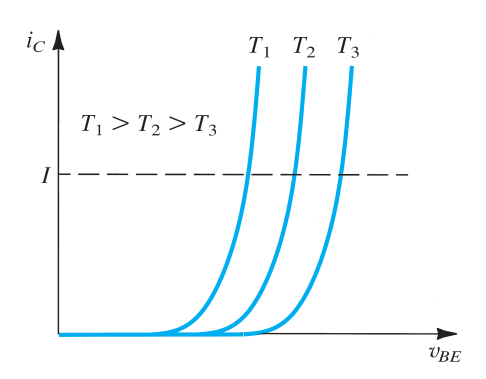


Figure 6.17 Effect of temperature on the i_C–v_{BE} characteristic. At a constant emitter current (broken line), v_{BE} changes by -2 mV/°C.

As in silicon diodes, the voltage across the emitter–base junction decreases by about 2 mV for each rise of 1°C in temperature, provided the junction is operating at a constant current. Figure 6.17 illustrates this temperature dependence by depicting i_C–v_{BE} curves for an *npn* transistor at three different temperatures.

EXERCISE

6.15 Consider a *pnp* transistor with $v_{EB} = 0.7$ V at $i_E = 1$ mA. Let the base be grounded, the emitter be fed by a 2-mA constant-current source, and the collector be connected to a –5-V supply through a 1-kΩ resistance. If the temperature increases by 30°C, find the changes in emitter and collector voltages. Neglect the effect of I_{CBO}.
Ans. –60 mV; 0 V

6.2.3 Dependence of i_C on the Collector Voltage—The Early Effect

When operated in the active region, practical BJTs show some dependence of the collector current on the collector voltage, with the result that, unlike the graph shown in Fig. 6.8, their i_C–v_{CB} characteristics are not perfectly horizontal straight lines. To see this dependence more

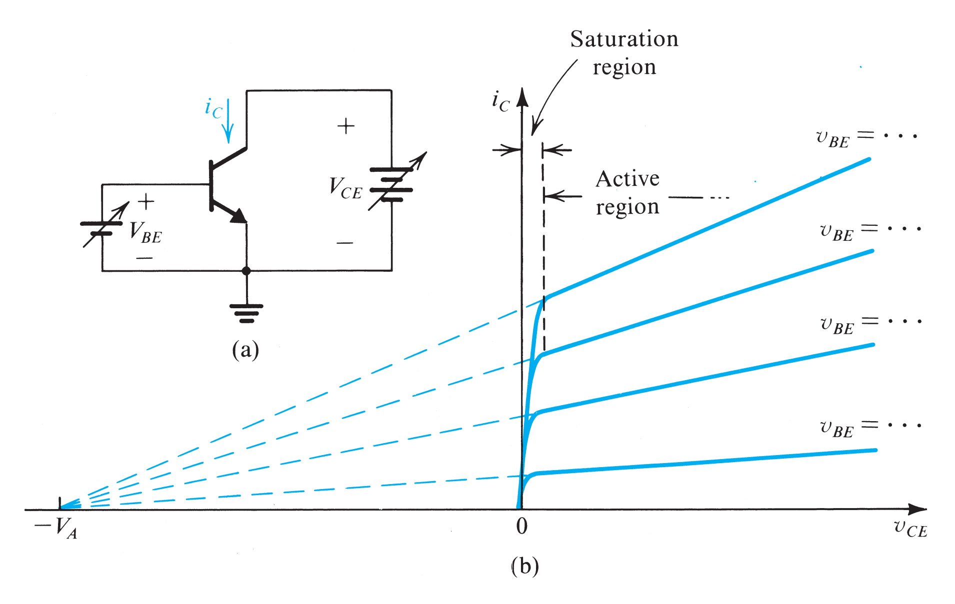


Figure 6.18 (**a**) Conceptual circuit for measuring the i_C–v_{CE} characteristics of the BJT. (**b**) The i_C–v_{CE} characteristics of a practical BJT.

clearly, consider the conceptual circuit shown in Fig. 6.18(a). The transistor is connected in the **common-emitter configuration**; that is, here the emitter serves as a common terminal between the input and output ports. The voltage V_{BE} can be set to any desired value by adjusting the dc source connected between base and emitter. At each value of V_{BE}, the corresponding i_C–v_{CE} characteristic curve can be measured point by point by varying the dc source connected between collector and emitter and measuring the corresponding collector current. The result is the family of i_C–v_{CE} characteristic curves shown in Fig. 6.18(b) and known as **common-emitter characteristics**.

At low values of v_{CE} (lower than about 0.3 V), as the collector voltage goes below that of the base by more than 0.4 V, the collector–base junction becomes forward biased and the transistor leaves the active mode and enters the saturation mode. Shortly, we shall look at the details of the i_C–v_{CE} curves in the saturation region. At this time, however, we wish to examine the characteristic curves in the active region in detail. We observe that the characteristic curves, though still straight lines, have finite slope. In fact, when extrapolated, the characteristic lines meet at a point on the negative v_{CE} axis, at $v_{CE} = -V_A$. The voltage V_A, a positive number, is a parameter for the particular BJT, with typical values in the range of 10 V to 100 V. As noted earlier, it is called the **Early voltage**, after J. M. Early, the engineering scientist who first studied this phenomenon.

At a given value of v_{BE}, increasing v_{CE} increases the reverse-bias voltage on the collector–base junction, and thus increases the width of the depletion region of this junction (refer to Fig. 6.4). This in turn results in a decrease in the **effective base width** *W*. Recalling that I_S is inversely proportional to *W* (Eq. 6.13), we see that I_S will increase and that i_C increases proportionally. This is the Early effect. For obvious reasons, it is also known as the **base-width modulation effect**.[8]

[8] Recall that the MOSFET's counterpart is the channel-length modulation effect. These two effects are remarkably similar and have been assigned the same name, Early effect.

The linear dependence of i_C on v_{CE} can be explicitly accounted for by assuming that I_S remains constant and including the factor $(1 + v_{CE}/V_A)$in the equation for i_C as follows:

$$i_C = I_S e^{v_{BE}/V_T}\left(1 + \frac{v_{CE}}{V_A}\right) \tag{6.18}$$

The nonzero slope of the i_C–v_{CE} straight lines indicates that the **output resistance** looking into the collector is not infinite. Rather, it is finite and defined by

$$r_o \equiv \left[\left.\frac{\partial i_C}{\partial v_{CE}}\right|_{v_{BE}\ =\ \text{constant}}\right]^{-1} \tag{6.19}$$

Using Eq. (6.18) we can show that

$$r_o = \frac{V_A + V_{CE}}{I_C} \tag{6.20}$$

where I_C and V_{CE} are the coordinates of the point at which the BJT is operating on the particular i_C–v_{CE} curve (i.e., the curve obtained for v_{BE} equal to constant value V_{BE} at which Eq. (6.19) is evaluated). Alternatively, we can write

$$r_o = \frac{V_A}{I'_C} \tag{6.21}$$

where I'_C is the value of the collector current with the Early effect neglected; that is,

$$I'_C = I_S e^{V_{BE}/V_T} \tag{6.22}$$

It is rarely necessary to include the dependence of i_C on v_{CE} in dc bias design and analysis that is performed by hand. Such an effect, however, can be easily included in the SPICE simulation of circuit operation, which is frequently used to "fine-tune" pencil-and-paper analysis or design.

The finite output resistance r_o can have a significant effect on the gain of transistor amplifiers. This is particularly the case in integrated-circuit amplifiers, as will be shown in Chapter 8. Fortunately, there are many situations in which r_o can be included relatively easily in pencil-and-paper analysis.

The output resistance r_o can be included in the circuit model of the transistor.[9] This is illustrated in Fig. 6.19, where we show the two large-signal circuit models of a

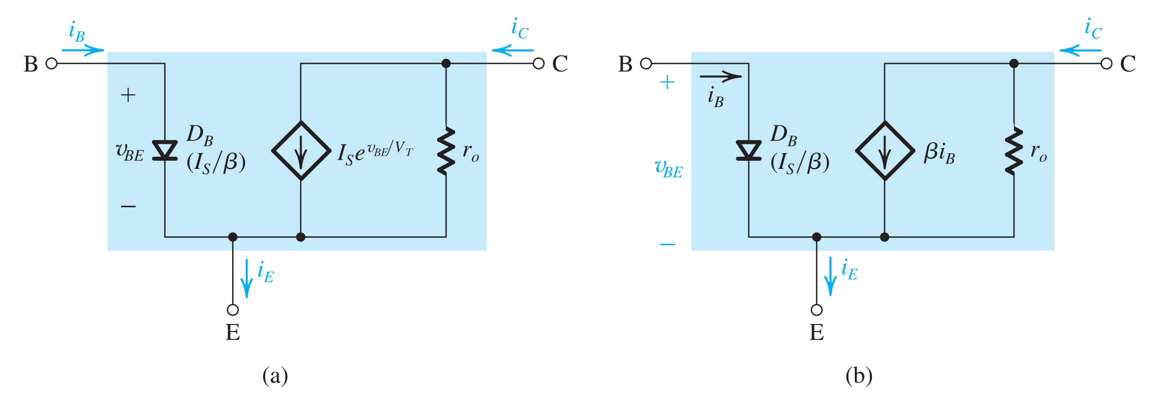


Figure 6.19 Large-signal, equivalent-circuit models of an *npn* BJT operating in the active mode in the common-emitter configuration with the output resistance r_o included.

[9] In applying Eq. (6.21) to determine r_o we will usually drop the prime and simply use $r_o = V_A/I_C$ where I_C is the collector current without the Early effect.

common-emitter *npn* transistor operating in the active mode, those in Fig 6.5(c) and (d), with the resistance r_o connected between the collector and the emitter terminals.

EXERCISES

6.16 Use the circuit model in Fig. 6.19(a) to express i_C in terms of e^{v_{BE}/V_T} and v_{CE} and thus show that this circuit is a direct representation of Eq. (6.18).

6.17 Find the output resistance of a BJT for which $V_A = 100$ V at $I_C = 0.1$, 1, and 10 mA.
Ans. 1 MΩ; 100 kΩ; 10 kΩ

6.18 Consider the circuit in Fig. 6.18(a). At $V_{CE} = 1$ V, V_{BE} is adjusted to yield a collector current of 1 mA. Then, while V_{BE} is kept constant, V_{CE} is raised to 11 V. Find the new value of I_C. For this transistor, $V_A = 100$ V.
Ans. 1.1 mA

6.2.4 An Alternative Form of the Common-Emitter Characteristics

An alternative way of expressing the transistor common-emitter characteristics is illustrated in Fig. 6.20. Here the base current i_B rather than the base–emitter voltage v_{BE} is used as a parameter. That is, each i_C–v_{CE} curve is measured with the base fed with a constant current I_B. The resulting characteristics, shown in Fig. 6.20(b), look similar to those in Fig. 6.18. Figure 6.20(c) shows an expanded view of the characteristics in the saturation region.

The Common-Emitter Current Gain β In the active region of the characteristics shown in Fig. 6.20(b) we have identified a particular point Q. Note that this operating point for the transistor is characterized by a base current I_B, a collector current I_C, and a collector–emitter voltage V_{CE}. The ratio I_C/I_B is the transistor β. However, there is another way to measure β: change the base current by an increment Δi_B and measure the resulting increment Δi_C, while keeping V_{CE} constant. This is illustrated in Fig. 6.20(b). The ratio $\Delta i_C/\Delta i_B$ should, according to our study thus far, yield an identical value for β. It turns out, however, that the latter value of β (called *incremental*, or ac, β) is a little different from the dc β (i.e., I_C/I_B). Such a distinction, however, is too subtle for our needs in this book. We shall use β to denote both dc and incremental values.[10]

The Saturation Voltage V_{CEsat} and Saturation Resistance R_{CEsat} Refer next to the expanded view of the common-emitter characteristics in the saturation region shown in Fig. 6.20(c). The "bunching together" of the curves in the saturation region implies that the incremental β is lower there than in the active region. A possible operating point in the saturation region is that labeled X. It is characterized by a base current I_B, a collector current I_{Csat}, and a collector–emitter voltage V_{CEsat}. From our previous discussion of saturation, recall that $I_{Csat} = \beta_{forced} I_B$, where $\beta_{forced} < \beta$.

[10]Manufacturers of bipolar transistors use h_{FE} to denote the dc value of β and h_{fe} to denote the incremental β. These symbols come from the h-parameter description of two-port networks (see Appendix C), with the subscript $F(f)$ denoting forward and $E(e)$ denoting common emitter.

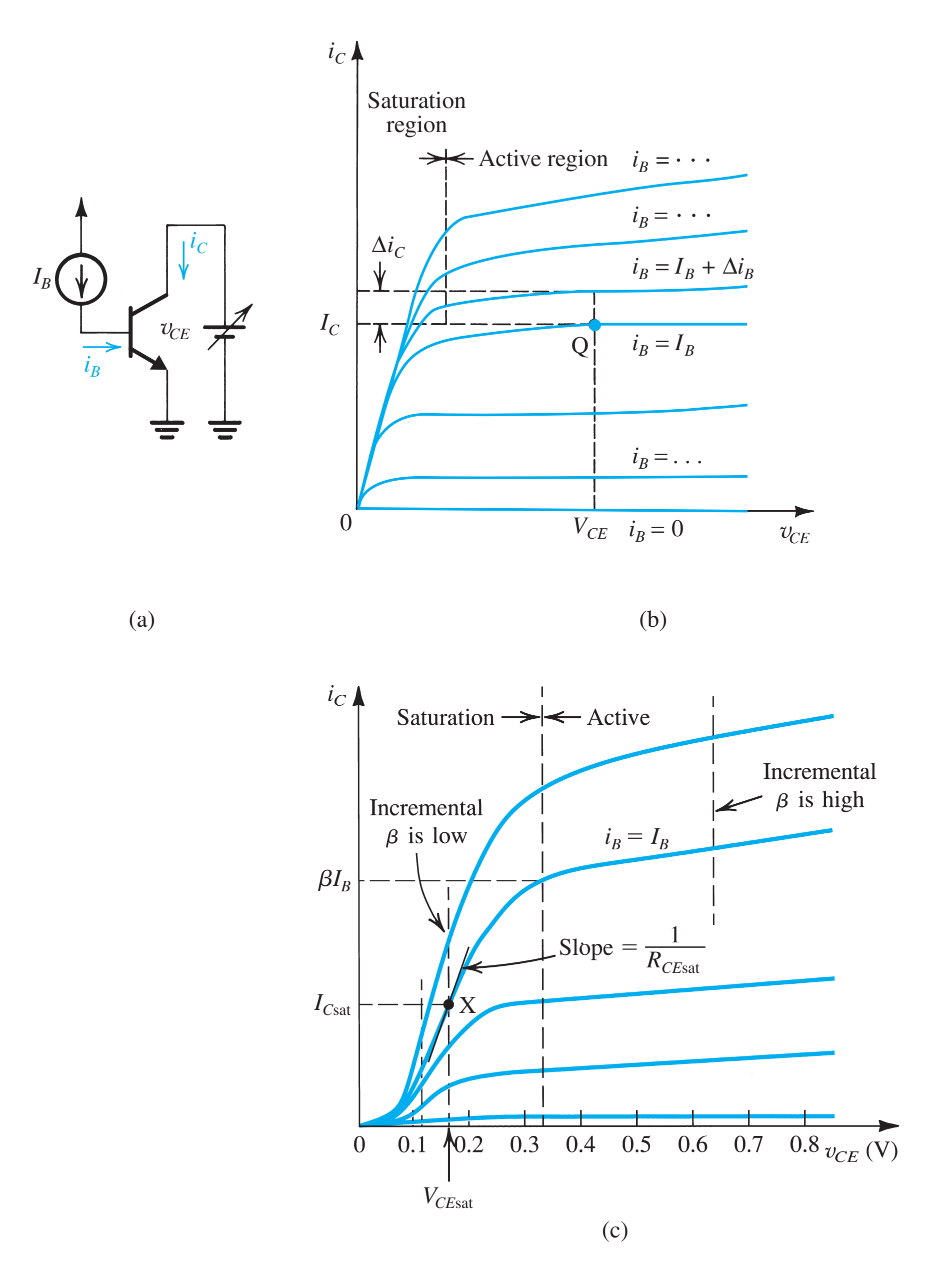


Figure 6.20 Common-emitter characteristics. (**a**) Basic CE circuit; note that in (**b**) the horizontal scale is expanded around the origin to show the saturation region in some detail. A much greater expansion of the saturation region is shown in (**c**) .

The i_C–v_{CE} curves in saturation are rather steep, indicating that the saturated transistor exhibits a low collector-to-emitter resistance R_{CEsat},

$$R_{CEsat} \equiv \left.\frac{\partial v_{CE}}{\partial i_C}\right|_{\substack{i_B = I_B \\ i_C = I_{Csat}}} \tag{6.23}$$

Typically, R_{CEsat} ranges from a few ohms to a few tens of ohms.

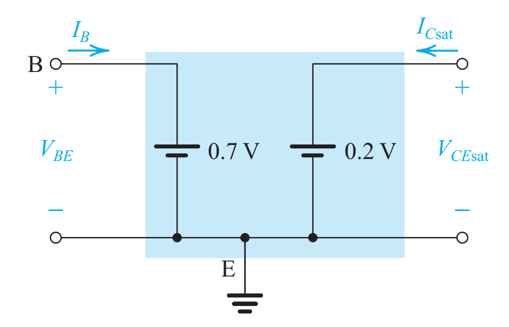


Figure 6.21 A simplified equivalent-circuit model of the saturated transistor.

That the collector-to-emitter resistance of a saturated BJT is small should have been anticipated from the fact that between C and E we now have two forward-conducting diodes in series[11] (see also Fig. 6.9).

A simple model for the saturated BJT is shown in Fig. 6.21. Here V_{BE} is assumed constant (approximately 0.7 V) and V_{CE} also is assumed constant, $V_{CEsat} \simeq 0.2$ V. That is, we have neglected the small saturation resistance R_{CEsat} for the sake of making the model simple for hand calculations.

Example 6.3

For the circuit in Fig. 6.22, it is required to determine the value of the voltage V_{BB} that results in the transistor operating

(a) in the active mode with $V_{CE} = 5$ V
(b) at the edge of saturation
(c) deep in saturation with $\beta_{forced} = 10$

For simplicity, assume that V_{BE} remains constant at 0.7 V. The transistor β is specified to be 50.

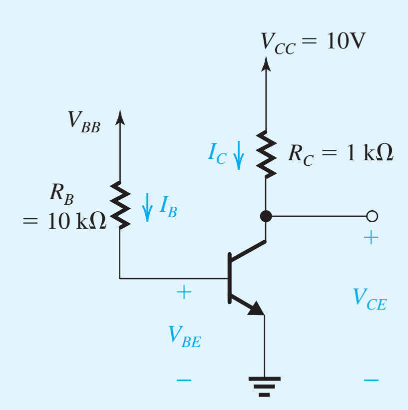


Figure 6.22 Circuit for Example 6.3.

[11]In the corresponding mode of operation for the MOSFET, the triode region, the resistance between drain and source is small because it is the resistance of the continuous (non-pinched-off) channel.

Example 6.3 *continued*

Solution

(a) To operate in the active mode with $V_{CE} = 5$ V,

$$I_C = \frac{V_{CC} - V_{CE}}{R_C}$$
$$= \frac{10-5}{1\text{ k}\Omega} = 5\text{ mA}$$
$$I_B = \frac{I_C}{\beta} = \frac{5}{50} = 0.1\text{ mA}$$

Now the required value of V_{BB} can be found as follows:

$$V_{BB} = I_B R_B + V_{BE}$$
$$= 0.1 \times 10 + 0.7 = 1.7\text{ V}$$

(b) Operation at the edge of saturation is obtained with $V_{CE} = 0.3$ V. Thus

$$I_C = \frac{10 - 0.3}{1} = 9.7\text{ mA}$$

Since, at the edge of saturation, I_C and I_B are still related by β,

$$I_B = \frac{9.7}{50} = 0.194\text{ mA}$$

The required value of V_{BB} can be determined as

$$V_{BB} = 0.194 \times 10 + 0.7 = 2.64\text{ V}$$

(c) To operate deep in saturation,

$$V_{CE} = V_{CE\text{sat}} \simeq 0.2\text{ V}$$

Thus,

$$I_C = \frac{10 - 0.2}{1} = 9.8\text{ mA}$$

We then use the value of forced β to determine the required value of I_B as

$$I_B = \frac{I_C}{\beta_{\text{forced}}} = \frac{9.8}{10} = 0.98\text{ mA}$$

and the required V_{BB} can now be found as

$$V_{BB} = 0.98 \times 10 + 0.7 = 10.5\text{ V}$$

Observe that once the transistor is in saturation, increasing V_{BB} and thus I_B results in negligible change in I_C since $V_{CE\text{sat}}$ will change only slightly. Thus I_C is said to *saturate*, which is the origin of the name "saturation mode of operation."

EXERCISES

6.19 Repeat Example 6.3 for $R_C = 10\ \text{k}\Omega$.
Ans. 0.8 V; 0.894 V; 1.68 V

6.20 For the circuit in Fig. 6.22, find V_{CE} for $V_{BB} = 0$ V.
Ans. + 10 V

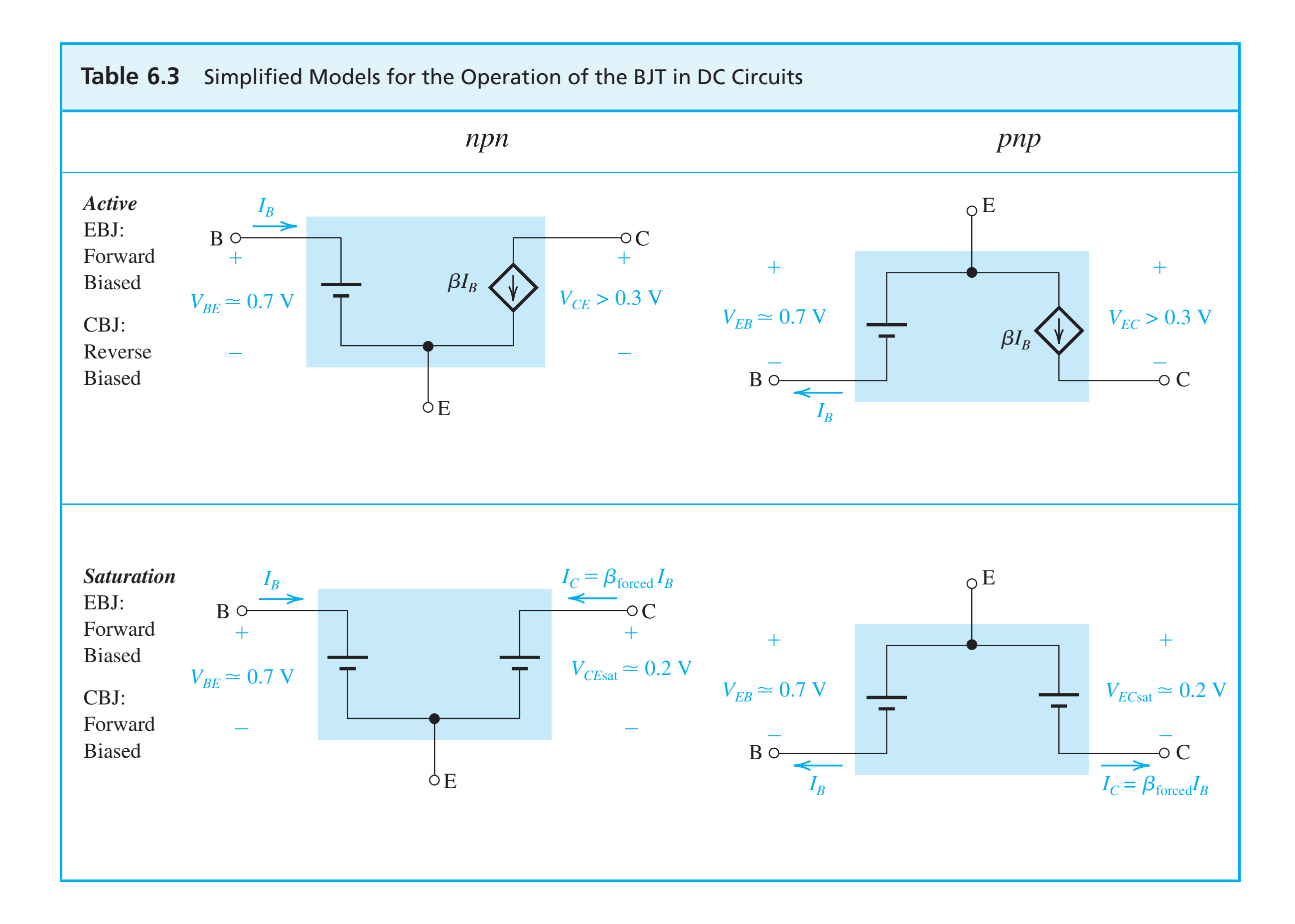
6.21 For the circuit in Fig. 6.22, let V_{BB} be set to the value obtained in Example 6.3, part (a), namely, $V_{BB} = 1.7$ V. Verify that the transistor is indeed operating in the active mode. Now, while keeping V_{BB} constant, find the value to which R_C should be increased in order to obtain (a) operation at the edge of saturation and (b) operation deep in saturation with $\beta_{\text{forced}} = 10$.
Ans. (a) 1.94 kΩ; (b) 9.8 kΩ

6.3 BJT Circuits at DC

We are now ready to consider the analysis of BJT circuits to which only dc voltages are applied. In the following examples we will use the simple model in which $|V_{BE}|$ of a conducting transistor is 0.7 V and $|V_{CE}|$ of a saturated transistor is 0.2 V, and we will neglect the Early effect. These models are shown in Table 6.3. Better models can, of course, be used to obtain more accurate results. This, however, is usually achieved at the expense of speed of analysis; more importantly, the attendant complexity could impede the circuit designer's ability to gain insight regarding circuit behavior. Accurate results using elaborate models can be obtained using circuit simulation with SPICE. This is almost always done in the final stages of a design and certainly before circuit fabrication. Computer simulation, however, is *not* a substitute for quick pencil-and-paper circuit analysis, an essential ability that aspiring circuit designers must master. The following series of examples is a step in that direction.

As will be seen, in analyzing a circuit the first question that one must answer is: *In which mode is the transistor operating*? In some cases, the answer will be obvious. For instance, a quick check of the terminal voltages will indicate whether the transistor is cut off or conducting. If it is conducting, we have to determine whether it is operating in the active mode or in saturation. In some cases, however, this may not be obvious. Needless to say, as the reader gains practice and experience in transistor circuit analysis and design, the answer will be apparent in a much larger proportion of problems. The answer, however, can always be determined by utilizing the following procedure.

Assume that the transistor is operating in the active mode and, using the active-mode model in Table 6.3, proceed to determine the various voltages and currents that correspond. Then check for consistency of the results with the assumption of active-mode operation; that is, is V_{CB} of an *npn* transistor greater than −0.4 V (or V_{CB} of a *pnp* transistor lower than 0.4 V)? If the answer is yes, then our task is complete. If the answer is no, assume saturation-mode operation and, using the saturation-mode model in Table 6.3, proceed to determine currents and voltages



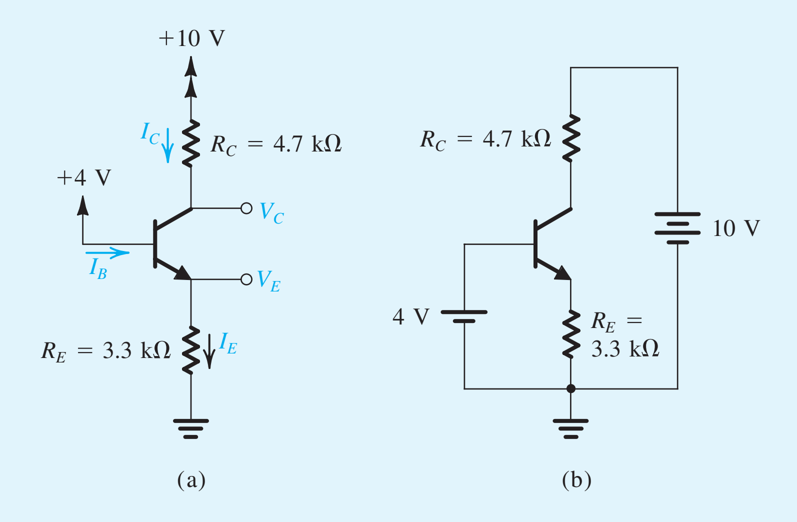
and then check for consistency of the results with the assumption of saturation-mode operation. Here the test is usually to compute the ratio I_C/I_B and to verify that it is lower than the transistor β (i.e., $\beta_{forced} < \beta$). Since β for a given transistor type varies over a wide range,[12] one must use the lowest specified β for this test. Finally, note that the order of these two assumptions can be reversed.

A Note on Units Except when otherwise specified, throughout this book we use a consistent set of units, namely, volts (V), milliamps (mA), and kilohms (kΩ).

[12] That is, if one buys BJTs of a certain part number, the manufacturer guarantees only that their values of β fall within a certain range, say 50 to 150.

Example 6.4

Consider the circuit shown in Fig. 6.23(a), which is redrawn in Fig. 6.23(b) to remind the reader of the convention employed throughout this book for indicating connections to dc sources. We wish to analyze this circuit to determine all node voltages and branch currents. We will assume that β is specified to be 100.



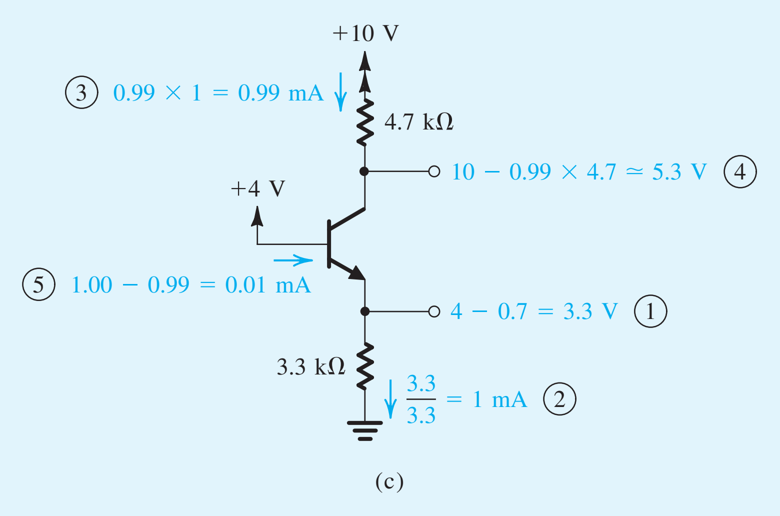


Figure 6.23 Analysis of the circuit for Example 6.4: (**a**) circuit; (**b**) circuit redrawn to remind the reader of the convention used in this book to show connections to the dc sources; (**c**) analysis with the steps numbered.

Example 6.4 *continued*

Solution

Glancing at the circuit in Fig. 6.23(a), we note that the base is connected to +4 V and the emitter is connected to ground through a resistance R_E. Therefore, it is reasonable to conclude that the base–emitter junction will be forward biased. Assuming that this is the case and assuming that V_{BE} is approximately 0.7 V, it follows that the emitter voltage will be

$$V_E = 4 - V_{BE} \simeq 4 - 0.7 = 3.3 \text{ V}$$

We are now in an opportune position; we know the voltages at the two ends of R_E and thus can determine the current I_E through it,

$$I_E = \frac{V_E - 0}{R_E} = \frac{3.3}{3.3} = 1 \text{ mA}$$

Since the collector is connected through R_C to the +10-V power supply, it appears possible that the collector voltage will be higher than the base voltage, which implies active-mode operation. Assuming that this is the case, we can evaluate the collector current from

$$I_C = \alpha I_E$$

The value of α is obtained from

$$\alpha = \frac{\beta}{\beta + 1} = \frac{100}{101} \simeq 0.99$$

Thus I_C will be given by

$$I_C = 0.99 \times 1 = 0.99 \text{ mA}$$

We are now in a position to use Ohm's law to determine the collector voltage V_C,

$$V_C = 10 - I_C R_C = 10 - 0.99 \times 4.7 \simeq +5.3 \text{ V}$$

Since the base is at +4 V, the collector–base junction is reverse biased by 1.3 V, and the transistor is indeed in the active mode as assumed.

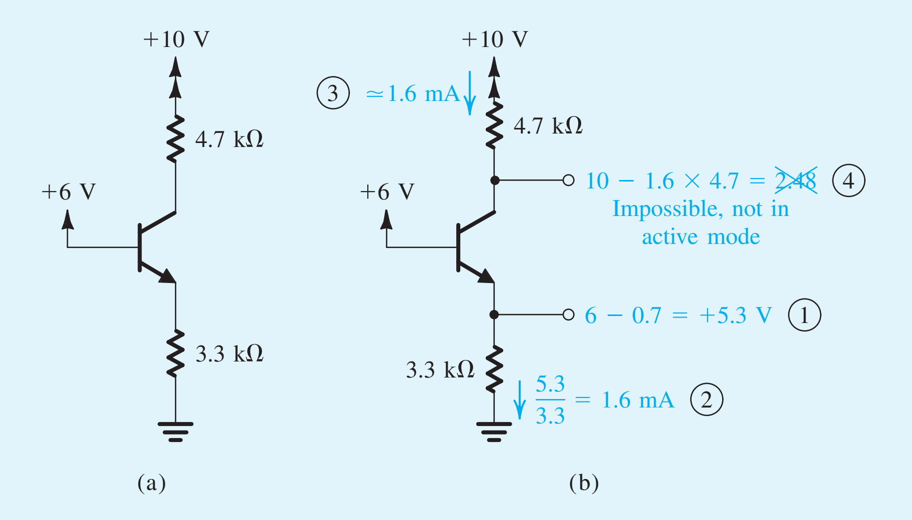
It remains only to determine the base current I_B, as follows:

$$I_B = \frac{I_E}{\beta + 1} = \frac{1}{101} \simeq 0.01 \text{ mA}$$

Before leaving this example, we wish to emphasize strongly the value of carrying out the analysis directly on the circuit diagram. Only in this way will one be able to analyze complex circuits in a reasonable length of time. Figure 6.23(c) illustrates the above analysis on the circuit diagram, with the order of the analysis steps indicated by the circled numbers.

Example 6.5

We wish to analyze the circuit of Fig. 6.24(a) to determine the voltages at all nodes and the currents through all branches. Note that this circuit is identical to that of Fig. 6.23 except that the voltage at the base is now +6 V. Assume that the transistor β is specified to be *at least* 50.



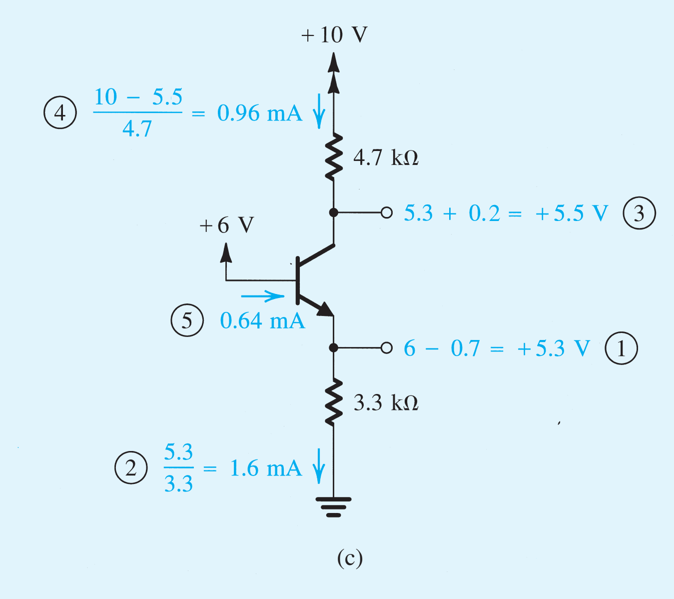


Figure 6.24 Analysis of the circuit for Example 6.5. Note that the circled numbers indicate the order of the analysis steps.

Example 6.5 *continued*

Solution

With +6 V at the base, the base–emitter junction will be forward biased; thus,

$$V_E = +6 - V_{BE} \simeq 6 - 0.7 = 5.3 \text{ V}$$

and

$$I_E = \frac{5.3}{3.3} = 1.6 \text{ mA}$$

Now, assuming active-mode operation, $I_C = \alpha I_E \simeq I_E$; thus,

$$V_C = +10 - 4.7 \times I_C \simeq 10 - 7.52 = 2.48 \text{ V}$$

The details of the analysis performed above are illustrated in Fig. 6.24(b).

Since the collector voltage calculated is less than the base voltage by 3.52 V, it follows that our original assumption of active-mode operation is incorrect. In fact, the transistor has to be in the *saturation* mode. Assuming this to be the case, the values of V_E and I_E will remain unchanged. The collector voltage, however, becomes

$$V_C = V_E + V_{CE\text{sat}} \simeq 5.3 + 0.2 = +5.5 \text{ V}$$

from which we can determine I_C as

$$I_C = \frac{10 - 5.5}{4.7} = 0.96 \text{ mA}$$

and I_B can now be found as

$$I_B = I_E - I_C = 1.6 - 0.96 = 0.64 \text{ mA}$$

Thus the transistor is operating at a forced β of

$$\beta_{\text{forced}} = \frac{I_C}{I_B} = \frac{0.96}{0.64} = 1.5$$

Since β_{forced} is less than the *minimum* specified value of β, the transistor is indeed saturated. We should emphasize here that in testing for saturation the minimum value of β should be used. By the same token, if we are designing a circuit in which a transistor is to be saturated, the design should be based on the minimum specified β. Obviously, if a transistor with this minimum β is saturated, then transistors with higher values of β will also be saturated. The details of the analysis are shown in Fig. 6.24(c), where the order of the steps used is indicated by the circled numbers.

Example 6.6

We wish to analyze the circuit in Fig. 6.25(a) to determine the voltages at all nodes and the currents through all branches. Note that this circuit is identical to that considered in Examples 6.4 and 6.5 except that now the base voltage is zero.

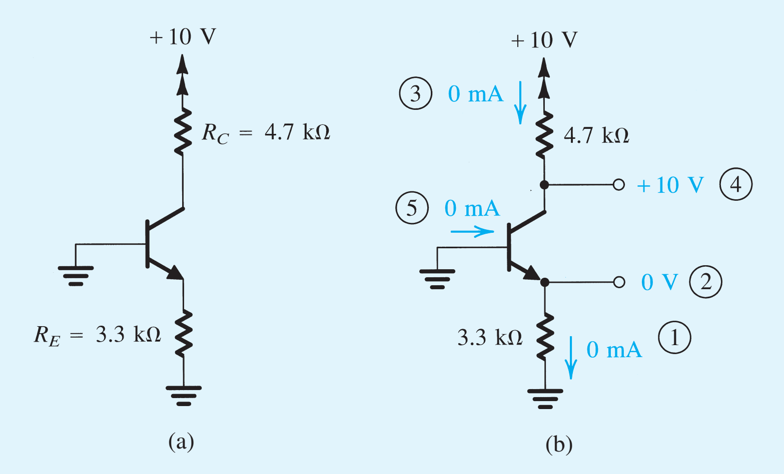


Figure 6.25 Example 6.6: (**a**) circuit; (**b**) analysis, with the order of the analysis steps indicated by circled numbers.

Solution

Since the base is at zero volts and the emitter is connected to ground through R_E, the base–emitter junction cannot conduct and the emitter current is zero. Also, the collector–base junction cannot conduct, since the *n*-type collector is connected through R_C to the positive power supply while the *p*-type base is at ground. It follows that the collector current will be zero. The base current will also have to be zero, and the transistor is in the *cutoff* mode of operation.

The emitter voltage will be zero, while the collector voltage will be equal to +10 V, since the voltage drops across R_E and R_C are zero. Figure 6.25(b) shows the analysis details.

EXERCISES

D6.22 For the circuit in Fig. 6.23(a), find the highest voltage to which the base can be raised while the transistor remains in the active mode. Assume $\alpha \simeq 1$.
Ans. +4.7 V

D6.23 Redesign the circuit of Fig. 6.23(a) (i.e., find new values for R_E and R_C) to establish a collector current of 0.5 mA and a reverse-bias voltage on the collector–base junction of 2 V. Assume $\alpha \simeq 1$.
Ans. $R_E = 6.6$ kΩ; $R_C = 8$ kΩ

D6.24 For the circuit in Fig. 6.24(a), find the value to which the base voltage should be changed so that the transistor operates in saturation with a forced β of 5.

Ans. +5.18 V

Example 6.7

We want to analyze the circuit of Fig. 6.26(a) to determine the voltages at all nodes and the currents through all branches.

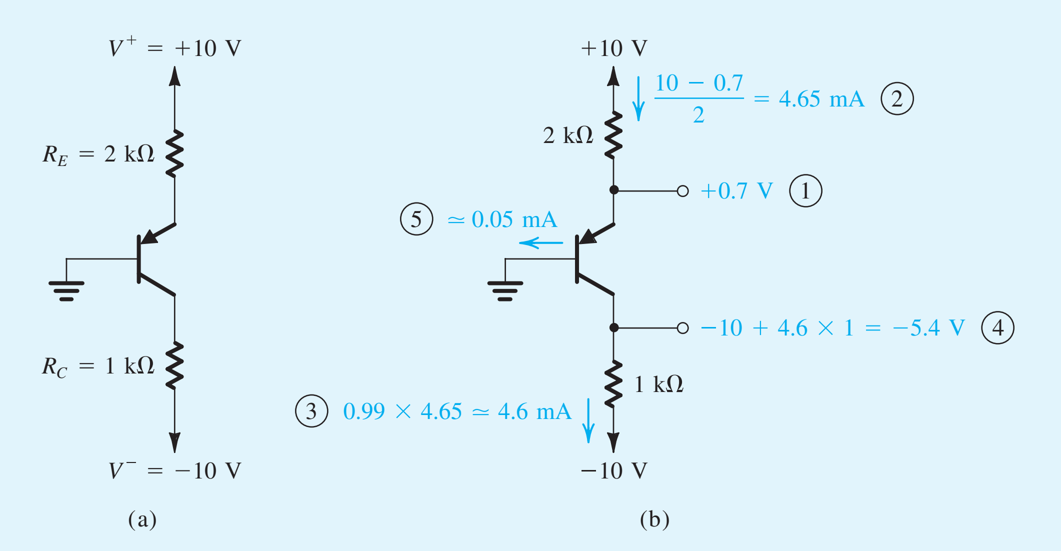


Figure 6.26 Example 6.7: (**a**) circuit; (**b**) analysis, with the steps indicated by circled numbers.

Solution

The base of this *pnp* transistor is grounded, while the emitter is connected to a positive supply ($V^+ = +10$ V) through R_E. It follows that the emitter–base junction will be forward biased with

$$V_E = V_{EB} \simeq 0.7 \text{ V}$$

Thus the emitter current will be given by

$$I_E = \frac{V^+ - V_E}{R_E} = \frac{10 - 0.7}{2} = 4.65 \text{ mA}$$

Since the collector is connected to a negative supply (more negative than the base voltage) through R_C, it is *possible* that this transistor is operating in the active mode. Assuming this to be the case, we obtain

$$I_C = \alpha I_E$$

Since no value for β has been given, we shall assume $\beta = 100$, which results in $\alpha = 0.99$. Since large variations in β result in small differences in α, this assumption will not be critical as far as determining the value of I_C is concerned. Thus,

$$I_C = 0.99 \times 4.65 = 4.6 \text{ mA}$$

The collector voltage will be

$$\begin{aligned} V_C &= V^- + I_C R_C \\ &= -10 + 4.6 \times 1 = -5.4 \text{ V} \end{aligned}$$

Thus the collector–base junction is reverse biased by 5.4 V, and the transistor is indeed in the active mode, which supports our original assumption.

It remains only to calculate the base current,

$$I_B = \frac{I_E}{\beta + 1} = \frac{4.65}{101} \simeq 0.05 \text{ mA}$$

Obviously, the value of β critically affects the base current. Note, however, that in this circuit the value of β will have no effect on the mode of operation of the transistor. Since β is generally an ill-specified parameter, this circuit represents a good design. As a rule, one should strive to *design the circuit such that its performance is as insensitive to the value of β as possible.* The analysis details are illustrated in Fig. 6.26(b).

EXERCISES

D6.25 For the circuit in Fig. 6.26(a), find the largest value to which R_C can be raised while the transistor remains in the active mode.
Ans. 2.26 kΩ

D6.26 Redesign the circuit of Fig. 6.26(a) (i.e., find new values for R_E and R_C) to establish a collector current of 1 mA and a reverse bias on the collector–base junction of 4 V. Assume $\alpha \simeq 1$.
Ans. $R_E = 9.3$ kΩ; $R_C = 6$ kΩ

Example 6.8

We want to analyze the circuit in Fig. 6.27(a) to determine the voltages at all nodes and the currents in all branches. Assume $\beta = 100$.

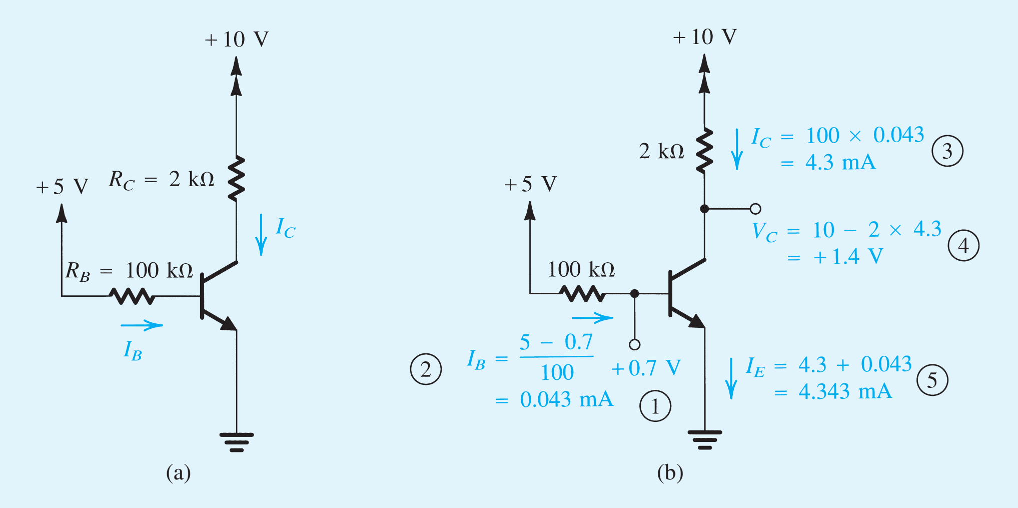


Figure 6.27 Example 6.8: (**a**) circuit; (**b**) analysis, with the steps indicated by the circled numbers.

Solution

The base–emitter junction is clearly forward biased. Thus,

$$I_B = \frac{+5 - V_{BE}}{R_B} \simeq \frac{5 - 0.7}{100} = 0.043 \text{ mA}$$

Assume that the transistor is operating in the active mode. We now can write

$$I_C = \beta I_B = 100 \times 0.043 = 4.3 \text{ mA}$$

The collector voltage can now be determined as

$$V_C = 10 - I_C R_C = 10 - 4.3 \times 2 = +1.4 \text{ V}$$

Since the base voltage V_B is

$$V_B = V_{BE} \simeq +0.7 \text{ V}$$

it follows that the collector–base junction is reverse biased by 0.7 V and the transistor is indeed in the active mode. The emitter current will be given by

$$I_E = (\beta + 1)I_B = 101 \times 0.043 \simeq 4.3 \text{ mA}$$

We note from this example that the collector and emitter currents depend critically on the value of β. In fact, if β were 10% higher, the transistor would leave the active mode and enter saturation. Therefore this clearly is a *bad* design. The analysis details are illustrated in Fig. 6.27(b).

EXERCISE

D6.27 The circuit of Fig. 6.27(a) is to be fabricated using a transistor type whose β is specified to be in the range of 50 to 150. That is, individual units of this same transistor type can have β values anywhere in this range. Redesign the circuit by selecting a new value for R_C so that all fabricated circuits are guaranteed to be in the active mode. What is the range of collector voltages that the fabricated circuits may exhibit?

Ans. $R_C = 1.5$ kΩ; $V_C = 0.3$ V to 6.8 V

Example 6.9

We want to analyze the circuit of Fig. 6.28(a) to determine the voltages at all nodes and the currents through all branches. The minimum value of β is specified to be 30.

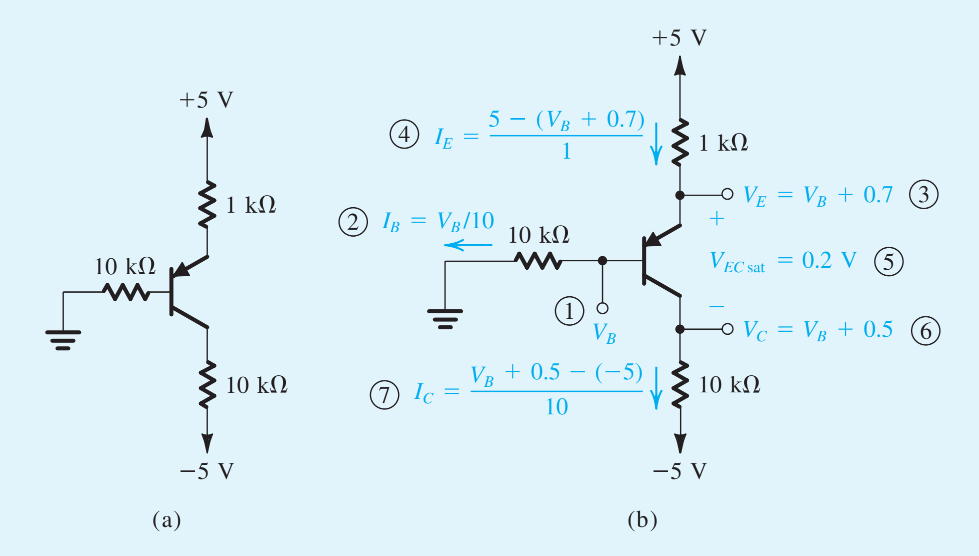


Figure 6.28 Example 6.9: (**a**) circuit; (**b**) analysis with steps numbered.

Solution

A quick glance at this circuit reveals that the transistor will be either active or saturated. Assuming active-mode operation and neglecting the base current, we see that the base voltage will be approximately zero volts, the emitter voltage will be approximately +0.7 V, and the emitter current will be approximately 4.3 mA. Since the maximum current that the collector can support while the transistor remains in the active mode is approximately 0.5 mA, it follows that the transistor is definitely saturated.

Example 6.9 *continued*

Assuming that the transistor is saturated and denoting the voltage at the base by V_B (refer to Fig. 6.28b), it follows that

$$V_E = V_B + V_{EB} \simeq V_B + 0.7$$
$$V_C = V_E - V_{ECsat} \simeq V_B + 0.7 - 0.2 = V_B + 0.5$$
$$I_E = \frac{+5 - V_E}{1} = \frac{5 - V_B - 0.7}{1} = 4.3 - V_B \quad \text{mA}$$
$$I_B = \frac{V_B}{10} = 0.1V_B \quad \text{mA}$$
$$I_C = \frac{V_C - (-5)}{10} = \frac{V_B + 0.5 + 5}{10} = 0.1V_B + 0.55 \quad \text{mA}$$

Using the relationship $I_E = I_B + I_C$, we obtain

$$4.3 - V_B = 0.1V_B + 0.1V_B + 0.55$$

which results in

$$V_B = \frac{3.75}{1.2} \simeq 3.13 \text{ V}$$

Substituting in the equations above, we obtain

$$V_E = 3.83 \text{ V}$$
$$V_C = 3.63 \text{ V}$$
$$I_E = 1.17 \text{ mA}$$
$$I_C = 0.86 \text{ mA}$$
$$I_B = 0.31 \text{ mA}$$

from which we see that the transistor is saturated, since the value of forced β is

$$\beta_{\text{forced}} = \frac{0.86}{0.31} \simeq 2.8$$

which is much smaller than the specified minimum β.

Example 6.10

We want to analyze the circuit of Fig. 6.29(a) to determine the voltages at all nodes and the currents through all branches. Assume $\beta = 100$.

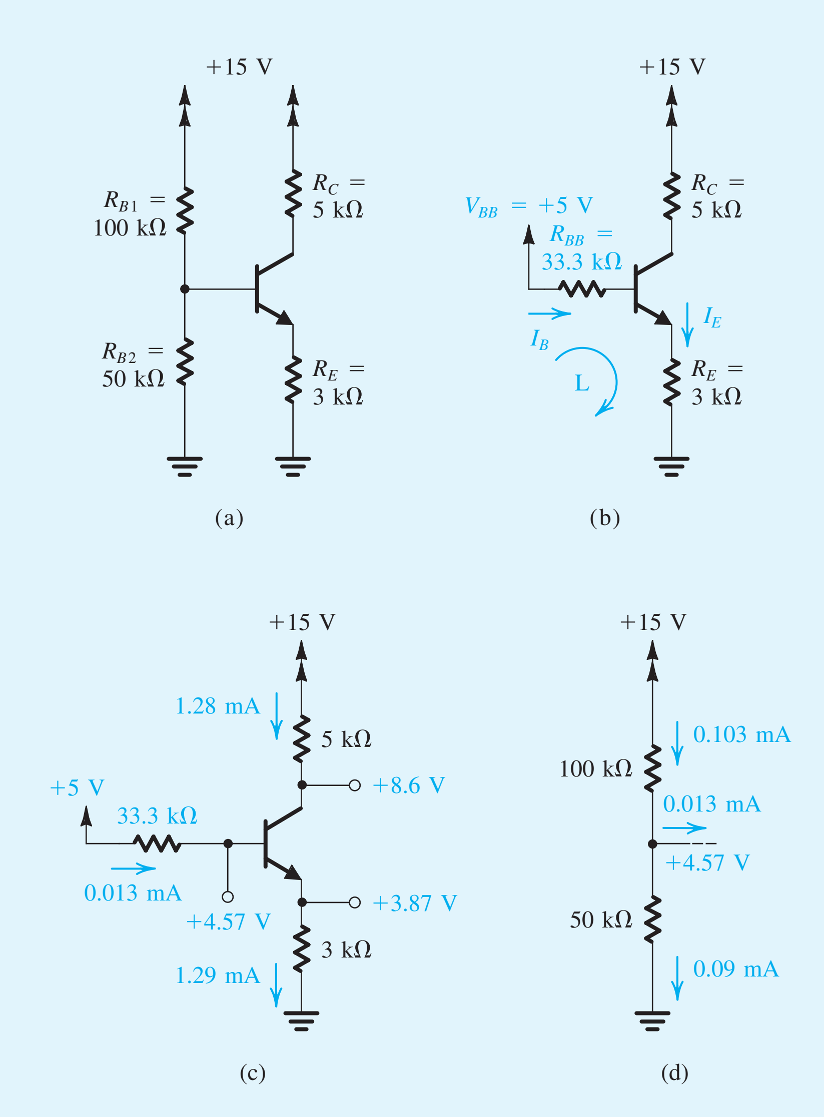


Figure 6.29 Circuits for Example 6.10.

Solution

The first step in the analysis consists of simplifying the base circuit using Thévenin's theorem. The result is shown in Fig. 6.29(b), where

$$V_{BB} = +15\frac{R_{B2}}{R_{B1}+R_{B2}} = 15\frac{50}{100+50} = +5 \text{ V}$$

$$R_{BB} = R_{B1} \,\|\, R_{B2} = 100 \,\|\, 50 = 33.3 \text{ k}\Omega$$

Example 6.10 *continued*

To evaluate the base or the emitter current, we have to write a loop equation around the loop labeled L in Fig. 6.29(b). Note, however, that the current through R_{BB} is different from the current through R_E. The loop equation will be

$$V_{BB} = I_B R_{BB} + V_{BE} + I_E R_E$$

Now, assuming active-mode operation, we replace I_B with

$$I_B = \frac{I_E}{\beta + 1}$$

and rearrange the equation to obtain

$$I_E = \frac{V_{BB} - V_{BE}}{R_E + [R_{BB}/(\beta + 1)]}$$

For the numerical values given we have

$$I_E = \frac{5 - 0.7}{3 + (33.3/101)} = 1.29 \text{ mA}$$

The base current will be

$$I_B = \frac{1.29}{101} = 0.0128 \text{ mA}$$

The base voltage is given by

$$V_B = V_{BE} + I_E R_E$$
$$= 0.7 + 1.29 \times 3 = 4.57 \text{ V}$$

We can evaluate the collector current as

$$I_C = \alpha I_E = 0.99 \times 1.29 = 1.28 \text{ mA}$$

The collector voltage can now be evaluated as

$$V_C = +15 - I_C R_C = 15 - 1.28 \times 5 = 8.6 \text{ V}$$

It follows that the collector is higher in potential than the base by 4.03 V, which means that the transistor is in the active mode, as had been assumed. The results of the analysis are given in Fig. 6.29(c, d).

EXERCISE

6.28 If the transistor in the circuit of Fig. 6.29(a) is replaced with another having half the value of β (i.e., $\beta = 50$), find the new value of I_C, and express the change in I_C as a percentage.
Ans. $I_C = 1.15$ mA; –10%

Example 6.11

We wish to analyze the circuit in Fig. 6.30(a) to determine the voltages at all nodes and the currents through all branches.

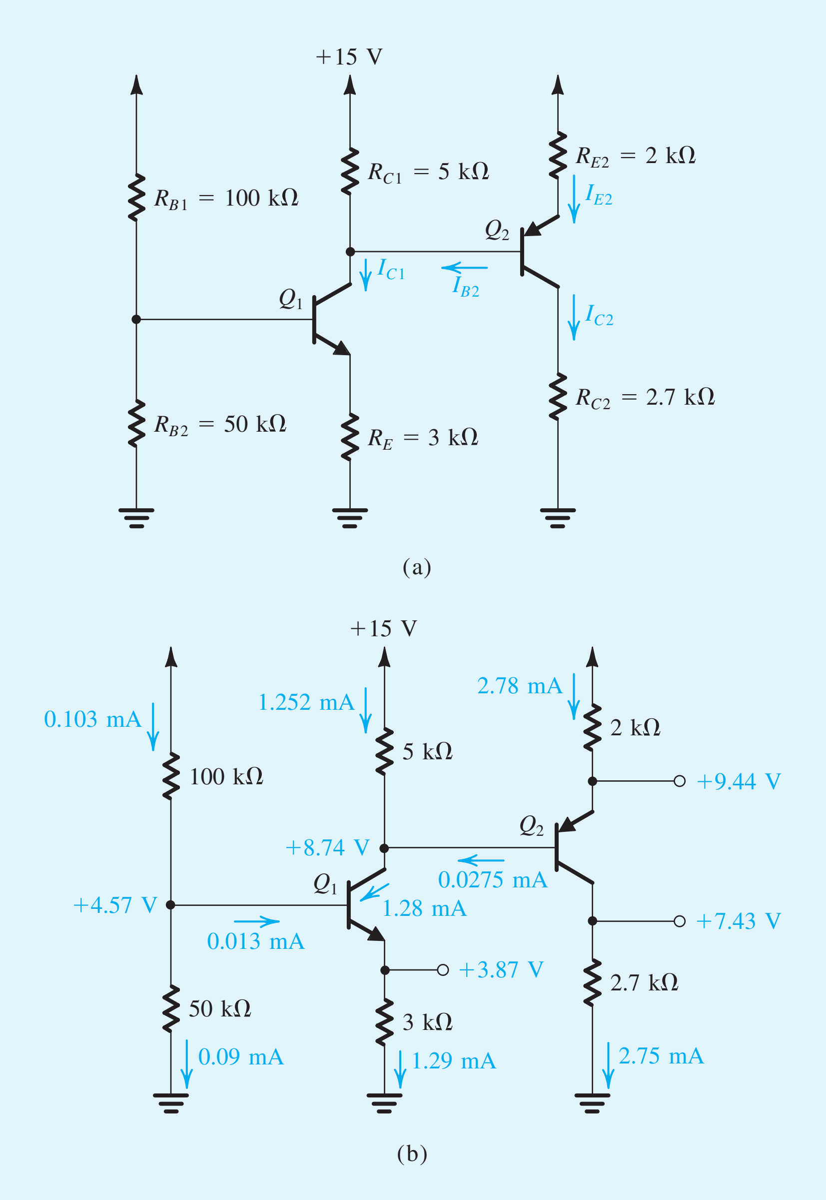


Figure 6.30 Circuits for Example 6.11.

Example 6.11 *continued*

Solution

We first recognize that part of this circuit is identical to the circuit we analyzed in Example 6.10—namely, the circuit of Fig. 6.29(a). The difference, of course, is that in the new circuit we have an additional transistor Q_2 together with its associated resistors R_{E2} and R_{C2}. Assume that Q_1 is still in the active mode. The following values will be identical to those obtained in the previous example:

$$V_{B1} = +4.57 \text{ V} \qquad I_{E1} = 1.29 \text{ mA}$$
$$I_{B1} = 0.0128 \text{ mA} \qquad I_{C1} = 1.28 \text{ mA}$$

However, the collector voltage will be different than previously calculated, since part of the collector current I_{C1} will flow in the base lead of Q_2 (I_{B2}). As a first approximation we may assume that I_{B2} is much smaller than I_{C1}; that is, we may assume that the current through R_{C1} is almost equal to I_{C1}. This will enable us to calculate V_{C1}:

$$\begin{aligned} V_{C1} &\simeq +15 - I_{C1}R_{C1} \\ &= 15 - 1.28 \times 5 = +8.6 \text{ V} \end{aligned}$$

Thus Q_1 is in the active mode, as had been assumed.

As far as Q_2 is concerned, we note that its emitter is connected to +15 V through R_{E2}. It is therefore safe to assume that the emitter–base junction of Q_2 will be forward biased. Thus the emitter of Q_2 will be at a voltage V_{E2} given by

$$V_{E2} = V_{C1} + V_{EB}\big|_{Q_2} \simeq 8.6 + 0.7 = +9.3 \text{ V}$$

The emitter current of Q_2 may now be calculated as

$$I_{E2} = \frac{+15 - V_{E2}}{R_{E2}} = \frac{15 - 9.3}{2} = 2.85 \text{ mA}$$

Since the collector of Q_2 is returned to ground via R_{C2}, it is possible that Q_2 is operating in the active mode. Assume this to be the case. We now find I_{C2} as

$$\begin{aligned} I_{C2} &= \alpha_2 I_{E2} \\ &= 0.99 \times 2.85 = 2.82 \text{ mA} \quad (\text{assuming } \beta_2 = 100) \end{aligned}$$

The collector voltage of Q_2 will be

$$V_{C2} = I_{C2}R_{C2} = 2.82 \times 2.7 = 7.62 \text{ V}$$

which is lower than V_{B2} by 0.98 V. Thus Q_2 is in the active mode, as assumed.

It is important at this stage to find the magnitude of the error incurred in our calculations by the assumption that I_{B2} is negligible. The value of I_{B2} is given by

$$I_{B2} = \frac{I_{E2}}{\beta_2 + 1} = \frac{2.85}{101} = 0.028 \text{ mA}$$

which is indeed much smaller than I_{C1} (1.28 mA). If desired, we can obtain more accurate results by iterating one more time, assuming I_{B2} to be 0.028 mA. The new values will be

$$\text{Current in } R_{C1} = I_{C1} - I_{B2} = 1.28 - 0.028 = 1.252 \text{ mA}$$

$$V_{C1} = 15 - 5 \times 1.252 = 8.74 \text{ V}$$

$$V_{E2} = 8.74 + 0.7 = 9.44 \text{ V}$$

$$I_{E2} = \frac{15 - 9.44}{2} = 2.78 \text{ mA}$$

$$I_{C2} = 0.99 \times 2.78 = 2.75 \text{ mA}$$

$$V_{C2} = 2.75 \times 2.7 = 7.43 \text{ V}$$

$$I_{B2} = \frac{2.78}{101} = 0.0275 \text{ mA}$$

Note that the new value of I_{B2} is very close to the value used in our iteration, and no further iterations are warranted. The final results are indicated in Fig. 6.30(b).

The reader justifiably might be wondering about the necessity for using an iterative scheme in solving a linear (or linearized) problem. Indeed, we can obtain the exact solution (if we can call anything we are doing with a first-order model exact!) by writing appropriate equations. The reader is encouraged to find this solution and then compare the results with those obtained above. It is important to emphasize, however, that in most such problems it is quite sufficient to obtain an approximate solution, provided we can obtain it quickly and, of course, correctly.

In the above examples, we frequently used a precise value of α to calculate the collector current. Since $\alpha \simeq 1$, the error in such calculations will be very small if one assumes $\alpha = 1$ and $I_C = I_E$. Therefore, except in calculations that depend critically on the value of α (e.g., the calculation of base current), one usually assumes $\alpha \simeq 1$.

EXERCISES

6.29 For the circuit in Fig. 6.30, find the total current drawn from the power supply. Hence find the power dissipated in the circuit.
Ans. 4.135 mA; 62 mW

6.30 The circuit in Fig. E6.30 is to be connected to the circuit in Fig. 6.30(a) as indicated; specifically, the base of Q_3 is to be connected to the collector of Q_2. If Q_3 has $\beta = 100$, find the new value of V_{C2} and the values of V_{E3} and I_{C3}.

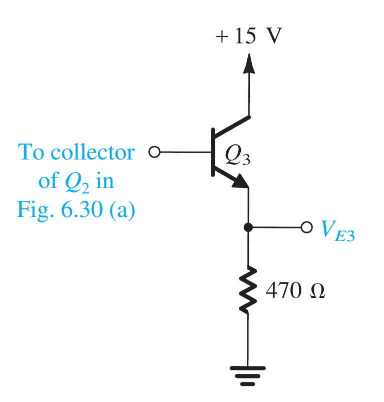


Figure E6.30

Ans. +7.06 V; +6.36 V; 13.4 mA

Example 6.12

We desire to evaluate the voltages at all nodes and the currents through all branches in the circuit of Fig. 6.31(a). Assume $\beta = 100$.

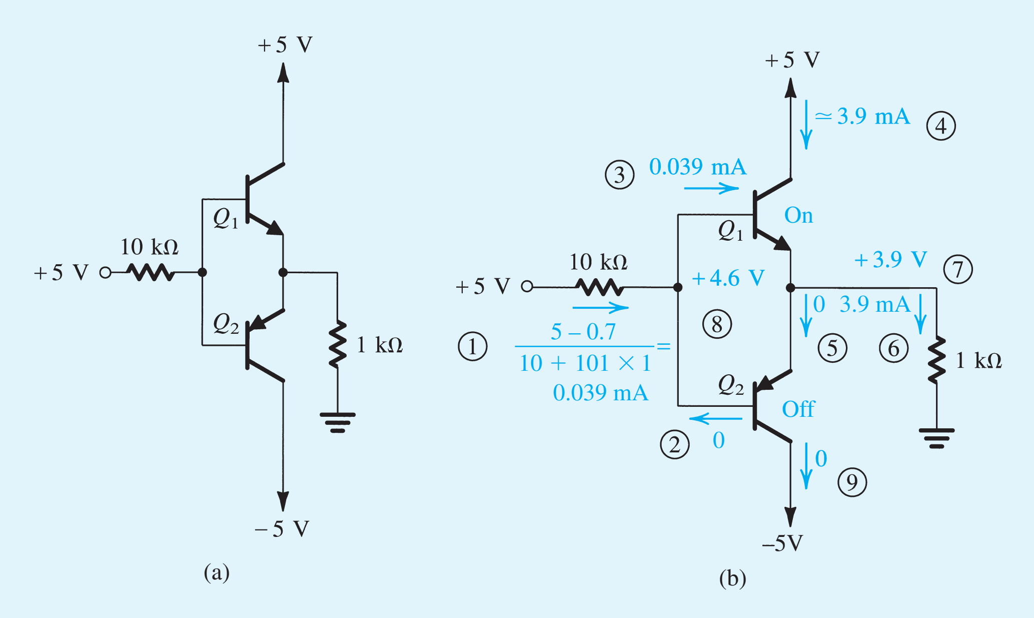


Figure 6.31 Example 6.12: (**a**) circuit; (**b**) analysis with the steps numbered.

Solution

By examining the circuit, we conclude that the two transistors Q_1 and Q_2 cannot be simultaneously conducting. Thus if Q_1 is on, Q_2 will be off, and vice versa. Assume that Q_2 is on. It follows that current will flow from ground through the 1-kΩ resistor into the emitter of Q_2. Thus the base of Q_2 will be at a negative voltage, and base current will be flowing out of the base through the 10-kΩ resistor and into the +5-V supply. This is impossible, since if the base is negative, current in the 10-kΩ resistor will have to flow into the base. Thus we conclude that our original assumption—that Q_2 is on—is incorrect. It follows that Q_2 will be off and Q_1 will be on.

The question now is whether Q_1 is active or saturated. The answer in this case is obvious: Since the base is fed with a +5-V supply and since base current flows into the base of Q_1, it follows that the base of Q_1 will be at a voltage lower than +5 V. Thus the collector–base junction of Q_1 is reverse biased and Q_1 is in the active mode. It remains only to determine the currents and voltages using techniques already described in detail. The results are given in Fig. 6.31(b).

EXERCISES

6.31 Solve the problem in Example 6.12 for the case of a voltage of –5 V feeding the bases. What voltage appears at the emitters?
Ans. –3.9 V

6.32 Solve the problem in Example 6.12 with the voltage feeding the bases changed to +10 V. Assume that $\beta_{\min} = 30$, and find V_E, V_B, I_{C1}, and I_{C2}.
Ans. +4.8 V; +5.5 V; 4.35 mA; 0

6.4 Transistor Breakdown and Temperature Effects



We conclude this chapter with a brief discussion of two important nonideal effects in the BJT: voltage breakdown, and the dependence of β on I_C and temperature.

6.4.1 Transistor Breakdown

The maximum voltages that can be applied to a BJT are limited by the EBJ and CBJ breakdown effects that follow the avalanche multiplication mechanism described in Section 3.5.3. Consider first the common-base configuration (Fig. 6.32(a)). The $i_C - v_{CB}$ characteristics in Fig. 6.32(b) indicate that for $i_E = 0$ (i.e., with the emitter open-circuited) the collector–base junction breaks down at a voltage denoted by BV_{CBO}. For $i_E > 0$, breakdown occurs at voltages smaller than BV_{CBO}. Typically, for discrete BJTs, BV_{CBO} is greater than 50 V.

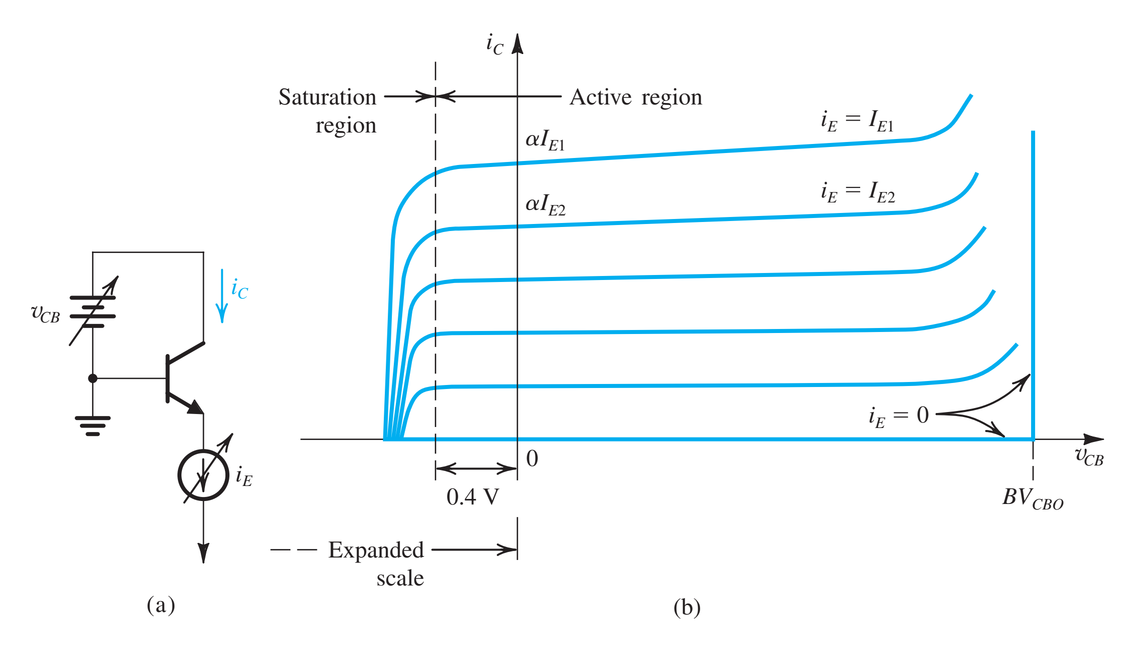


Figure 6.32 The BJT common-base characteristics including the transistor breakdown region.

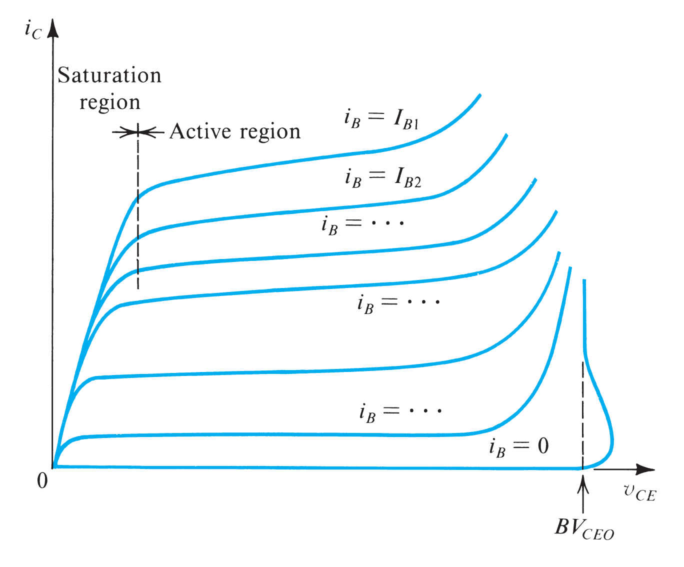
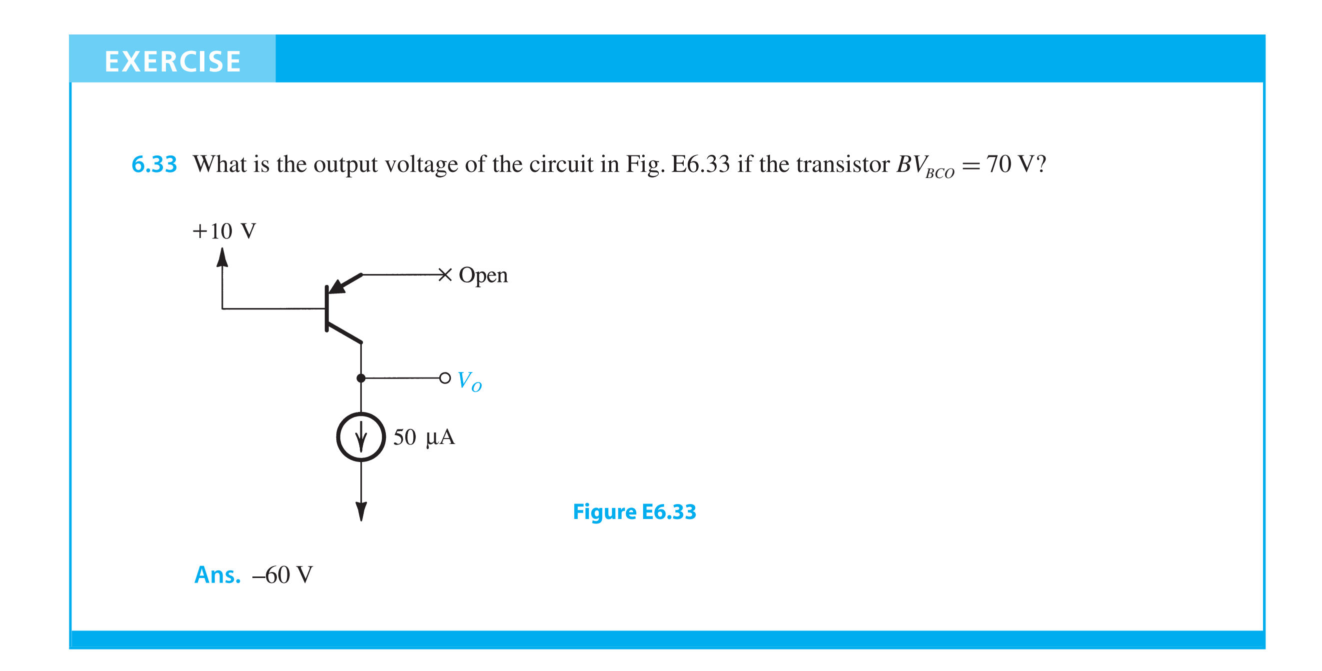


Figure 6.33 The BJT common-emitter characteristics including the breakdown region.

Next consider the common-emitter characteristics of Fig. 6.33, which show breakdown occurring at a voltage BV_{CEO}. Here, although breakdown is still of the avalanche type, the effects on the characteristics are more complex than in the common-base configuration. We will not explain these in detail; it is sufficient to point out that typically BV_{CEO} is about half BV_{CBO}. On transistor data sheets, BV_{CEO} is sometimes referred to as the **sustaining voltage** LV_{CEO}.

Breakdown of the CBJ in either the common-base or common-emitter configuration is not destructive as long as the power dissipation in the device is kept within safe limits. This, however, is not the case with the breakdown of the emitter–base junction. The EBJ breaks down in an avalanche manner at a voltage BV_{EBO} much smaller than BV_{CBO}. Typically, BV_{EBO} is in the range of 6 V to 8 V, and the breakdown is destructive in the sense that the β of the transistor is permanently reduced. This does not prevent use of the EBJ as a zener diode to generate reference voltages in IC design. In such applications one is not concerned with the β-degradation effect. A circuit arrangement to prevent EBJ breakdown in IC amplifiers will be discussed in Chapter 13. Transistor breakdown and the maximum allowable power dissipation are important parameters in the design of power amplifiers (Chapter 12).



EXERCISE

6.33 What is the output voltage of the circuit in Fig. E6.33 if the transistor $BV_{BCO} = 70$ V?

Figure E6.33

Ans. –60 V

6.4.2 Dependence of β on I_C and Temperature

Throughout this chapter we have assumed that the transistor common-emitter dc current gain, β or h_{FE}, is constant for a given transistor. In fact, β depends on the dc current at which the transistor is operating, as shown in Fig. 6.34. The physical processes that give rise to this dependence are beyond the scope of this book. Note, however, that there is a current range over which β is highest. Normally, one arranges to operate the transistor at a current within this range.

Figure 6.34 also shows the dependence of β on temperature. The fact that β increases with temperature can lead to serious problems in transistors that operate at large power levels (see Chapter 12).

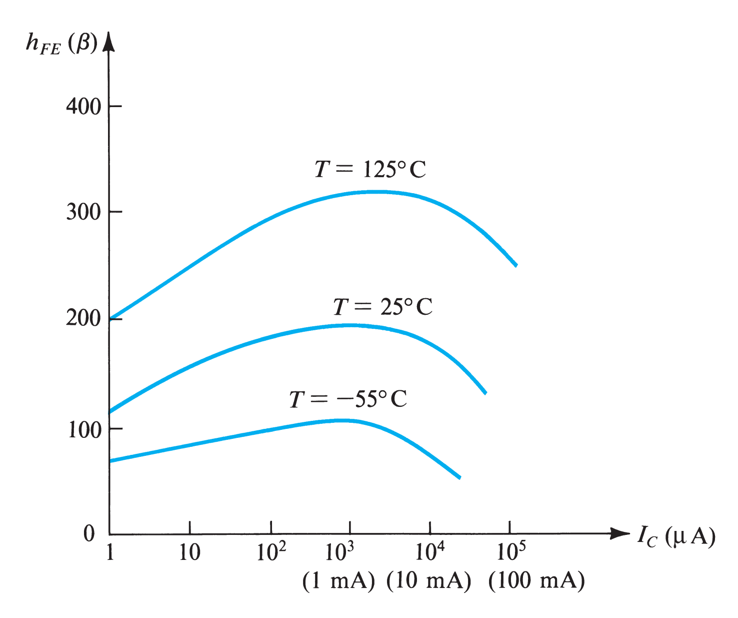


Figure 6.34 Typical dependence of β on I_C and on temperature in an integrated-circuit *npn* silicon transistor intended for operation around 1 mA.

Summary

- Depending on the bias conditions on its two junctions, the BJT can operate in one of three possible modes: cutoff (both junctions reverse biased), active (the EBJ forward biased and the CBJ reverse biased), and saturation (both junctions forward biased). Refer to Table 6.1.
- For amplifier applications, the BJT is operated in the active mode. Switching applications make use of the cutoff and saturation modes.
- A BJT operating in the active mode provides a collector current $i_C = I_S e^{|v_{BE}|/V_T}$. The base current $i_B = i_C/\beta$, and the emitter current $i_E = i_C + i_B$. Also, $i_C = \alpha i_E$, and thus $\beta = \alpha/(1-\alpha)$ and $\alpha = \beta/(\beta+1)$. See Table 6.2.
- To ensure operation in the active mode, the collector voltage of an *npn* transistor must be kept higher than approximately 0.4 V below the base voltage. For a *pnp* transistor, the collector voltage must be lower than approximately 0.4 V above the base voltage. Otherwise, the CBJ becomes forward biased, and the transistor enters the saturation region.
- At a constant collector current, the magnitude of the base–emitter voltage decreases by about 2 mV for every 1°C rise in temperature.
- The BJT will be at the edge of saturation when $|v_{CE}|$ is reduced to about 0.3 V. In saturation, $|v_{CE}| \simeq 0.2$ V, and the ratio of i_C to i_B is lower than β (i.e., $\beta_{\text{forced}} < \beta$).
- In the active mode, i_C shows a slight dependence on v_{CE}. This phenomenon, known as the Early effect, is modeled by ascribing a finite (i.e., noninfinite) output resistance to the BJT: $r_o = |V_A|/I'_C$, where V_A is the Early voltage and I'_C is the dc collector current without the Early effect taken into account. In discrete circuits, r_o plays a minor role and can usually be neglected. This is *not* the case, however, in integrated-circuit design (Chapter 8).
- The dc analysis of transistor circuits is greatly simplified by assuming that $|V_{BE}| \simeq 0.7$ V. Refer to Table 6.3.
- If the BJT is conducting, one assumes it is operating in the active mode and, using the active-mode model, proceeds to determine all currents and voltages. The validity of the initial assumption is then checked by determining whether the CBJ is reverse biased. If it is, the analysis is complete; otherwise, we assume the BJT is operating in saturation and redo the analysis, using the saturation-mode model and checking at the end that $I_C < \beta I_B$.

Computer Simulations Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as allowable signal swing and amplifier nonlinear distortion. Instructions to assist in setting up PSPice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 6.1: Device Structure and Physical Operation

6.1 The terminal voltages of various *npn* transistors are measured during operation in their respective circuits with the following results:

Case	E	B	C	Mode
1	0	0.7	0.7	
2	0	0.8	0.1	
3	−0.7	0	1.0	
4	−0.7	0	−0.6	
5	1.3	2.0	5.0	
6	0	0	5.0	

In this table, where the entries are in volts, 0 indicates the reference terminal to which the black (negative) probe of the voltmeter is connected. For each case, identify the mode of operation of the transistor.

6.2 Two transistors, fabricated with the same technology but having different junction areas, when operated at a base-emitter voltage of 0.75 V, have collector currents of 0.5 mA and 2 mA. Find I_S for each device. What are the relative junction areas?

6.3 In a particular technology, a small BJT operating at $v_{BE} = 30V_T$ conducts a collector current of 200 μA. What is the corresponding saturation current? For a transistor in the same technology but with an emitter junction that is 32 times larger, what is the saturation current? What current will this transistor conduct at $v_{BE} = 30V_T$? What is the base–emitter voltage of the latter transistor at $i_C = 1$ mA? Assume active-mode operation in all cases.

6.4 Two transistors have EBJ areas as follows: $A_{E1} = 200\ \mu\text{m} \times 200\ \mu\text{m}$ and $A_{E2} = 0.4\ \mu\text{m} \times 0.4\ \mu\text{m}$. If the two transistors are operated in the active mode and conduct equal collector currents, what do you expect the difference in their v_{BE} values to be?

6.5 Find the collector currents that you would expect for operation at $v_{BE} = 700$ mV for transistors for which $I_S = 10^{-13}$ A and $I_S = 10^{-18}$ A. For the transistor with the larger EBJ, what is the v_{BE} required to provide a collector current equal to that provided by the smaller transistor at $v_{BE} =$ 700 mV? Assume active-mode operation in all cases.

6.6 In this problem, we contrast two BJT integrated-circuit fabrication technologies: For the "old" technology, a typical *npn* transistor has $I_S = 2 \times 10^{-15}$ A, and for the "new" technology, a typical *npn* transistor has $I_S = 2 \times 10^{-18}$ A. These typical devices have vastly different junction areas and base width. For our purpose here we wish to determine the v_{BE} required to establish a collector current of 1 mA in each of the two typical devices. Assume active-mode operation.

6.7 Consider an *npn* transistor whose base–emitter drop is 0.76 V at a collector current of 5 mA. What current will it conduct at $v_{BE} = 0.70$ V? What is its base–emitter voltage for $i_C = 5$ μA?

6.8 In a particular BJT, the base current is 10 μA, and the collector current is 800 μA. Find β and α for this device.

6.9 Find the values of β that correspond to α values of 0.5, 0.8, 0.9, 0.95, 0.98, 0.99, 0.995, and 0.999.

6.10 Find the values of α that correspond to β values of 1, 2, 10, 20, 50, 100, 200, 500, and 1000.

***6.11** Show that for a transistor with α close to unity, if α changes by a small per-unit amount $(\Delta\alpha/\alpha)$, the corresponding per-unit change in β is given approximately by

$$\frac{\Delta\beta}{\beta} \simeq \beta\left(\frac{\Delta\alpha}{\alpha}\right)$$

Now, for a transistor whose nominal β is 100, find the percentage change in its α value corresponding to a drop in its β of 10%.

6.12 An *npn* transistor of a type whose β is specified to range from 50 to 300 is connected in a circuit with emitter grounded, collector at +10 V, and a current of 10 μA injected into the base. Calculate the range of collector and emitter currents that can result. What is the maximum power dissipated in the transistor? (*Note:* Perhaps you can see why this is a bad way to establish the operating current in the collector of a BJT.)

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

6.13 A BJT is specified to have $I_S = 5 \times 10^{-15}$ A and β that falls in the range of 50 to 200. If the transistor is operated in the active mode with v_{BE} set to 0.700 V, find the expected range of i_C, i_B, and i_E.

6.14 Measurements made on a number of transistors operating in the active mode with $i_E = 1$ mA indicate base currents of 10 μA, 20 μA, and 50 μA. For each device, find i_C, β, and α.

6.15 Measurements of V_{BE} and two terminal currents taken on a number of *npn* transistors operating in the active mode are tabulated below. For each, calculate the missing current value as well as α, β, and I_S as indicated by the table.

Transistor	a	b	c	d	e
V_{BE} (mV)	700	690	580	780	820
I_C (mA)	1.000	1.000		10.10	
I_B (μA)	10		5	120	1050
I_E (mA)		1.020	0.235		75.00
α					
β					
I_S					

6.16 When operated in the active mode, a particular *npn* BJT conducts a collector current of 1 mA and has $v_{BE} = 0.70$ V and $i_B = 10$ μA. Use these data to create specific transistor models of the form shown in Fig. 6.5(a) to (d).

6.17 Using the *npn* transistor model of Fig. 6.5(b), consider the case of a transistor for which the base is connected to ground, the collector is connected to a 5-V dc source through a 2-kΩ resistor, and a 2-mA current source is connected to the emitter with the polarity so that current is drawn out of the emitter terminal. If $\beta = 100$ and $I_S = 5 \times 10^{-15}$ A, find the voltages at the emitter and the collector and calculate the base current.

D 6.18 Consider an *npn* transistor operated in the active mode and represented by the model of Fig. 6.5(d). Let the transistor be connected as indicated by the equivalent circuit shown in Fig. 6.6(b). It is required to calculate the values of R_B and R_C that will establish a collector current I_C of 0.5 mA and a collector-to-emitter voltage V_{CE} of 1 V. The BJT is specified to have $\beta = 50$ and $I_S = 5 \times 10^{-15}$ A.

6.19 An *npn* transistor has a CBJ with an area 100 times that of the EBJ. If $I_S = 10^{-15}$ A, find the voltage drop across EBJ and across CBJ when each is forward biased and conducting a current of 1 mA. Also find the forward current each junction would conduct when forward biased with 0.5 V.

***6.20** We wish to investigate the operation of the *npn* transistor in saturation using the model of Fig. 6.9. Let $I_S = 10^{-15}$ A, $v_{BE} = 0.7$ V, $\beta = 100$, and $I_{SC}/I_S = 100$. For each of three values of v_{CE} (namely, 0.4 V, 0.3 V, and 0.2 V), find v_{BC}, i_{BC}, i_{BE}, i_B, i_C, and i_C/i_B. Present your results in tabular form. Also find v_{CE} that results in $i_C = 0$.

***6.21** Use Eqs. (6.14), (6.15), and (6.16) to show that an *npn* transistor operated in saturation exhibits a collector-to-emitter voltage, V_{CEsat}, given by

$$V_{CEsat} = V_T \ln\left[\left(\frac{I_{SC}}{I_S}\right)\frac{1+\beta_{forced}}{1-\beta_{forced}/\beta}\right]$$

Use this relationship to evaluate V_{CEsat} for $\beta_{forced} = 50$, 10, 5, and 1 for a transistor with $\beta = 100$ and with a CBJ area 100 times that of the EBJ. Present your results in a table.

6.22 Consider the *pnp* large-signal model of Fig. 6.11(b) applied to a transistor having $I_S = 10^{-14}$ A and $\beta = 50$. If the emitter is connected to ground, the base is connected to a current source that pulls 10 μA out of the base terminal, and the collector is connected to a negative supply of −5 V via a 8.2-kΩ resistor, find the collector voltage, the emitter current, and the base voltage.

6.23 A *pnp* transistor has $v_{EB} = 0.7$ V at a collector current of 1 mA. What do you expect v_{EB} to become at $i_C = 10$ mA? At $i_C = 100$ mA?

6.24 A *pnp* transistor modeled with the circuit in Fig. 6.11 (b) is connected with its base at ground, collector at –2.0 V, and a 1-mA current is injected into its emitter. If the transistor is said to have $\beta = 10$, what are its base and collector currents? In which direction do they flow? If $I_S = 10^{-15}$ A, what voltage results at the emitter? What does the collector current become if a transistor with $\beta = 1000$ is substituted? (*Note:* The fact that the collector current changes by less than 10% for a large change in β illustrates that this is a good way to establish a specific collector current.)

6.25 A *pnp* power transistor operates with an emitter-to-collector voltage of 5 V, an emitter current of 5 A, and $V_{EB} = 0.8$ V. For $\beta = 20$, what base current is required? What is I_S for this transistor? Compare the emitter–base junction area of this transistor with that of a small-signal

transistor that conducts $i_C = 1$ mA with $v_{EB} = 0.70$ V. How much larger is it?

6.26 While Fig. 6.5 provides four possible large-signal equivalent circuits for the *npn* transistor, only two equivalent circuits for the *pnp* transistor are provided in Fig. 6.11. Supply the missing two.

6.27 By analogy to the *npn* case shown in Fig. 6.9, give the equivalent circuit of a *pnp* transistor in saturation.

Section 6.2: Current–Voltage Characteristics

6.28 For the circuits in Fig. P6.28, assume that the transistors have very large β. Some measurements have been made on these circuits, with the results indicated in the figure. Find the values of the other labeled voltages and currents.

6.29 Measurements on the circuits of Fig. P6.29 produce labeled voltages as indicated. Find the value of β for each transistor.

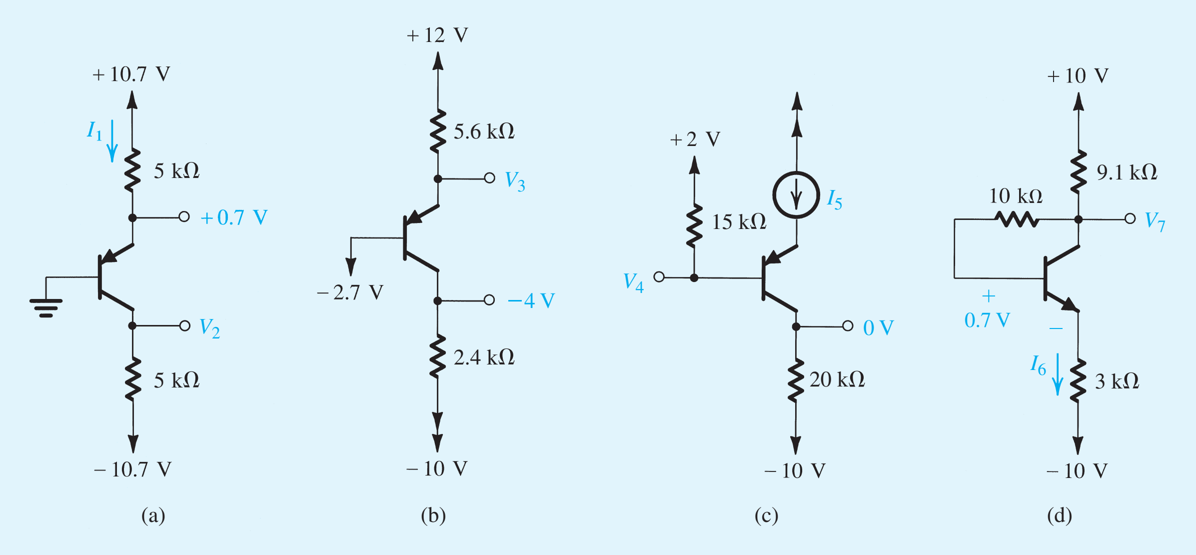


Figure P6.28

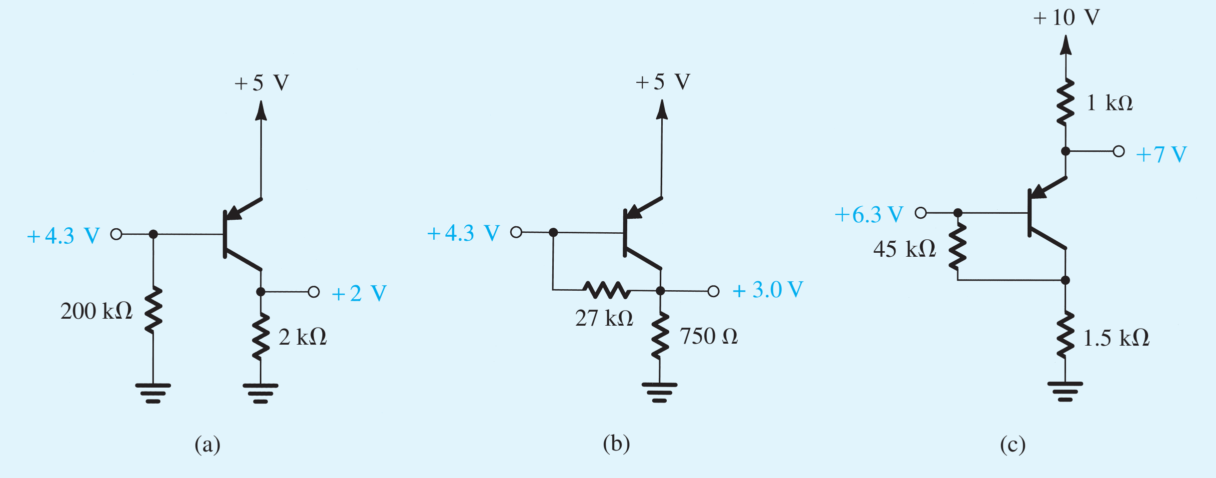


Figure P6.29

6.30 A very simple circuit for measuring β of an *npn* transistor is shown in Fig. P6.30. In a particular design, V_{CC} is provided by a 9-V battery; M is a current meter with a 50-μA full scale and relatively low resistance that you can neglect for our purposes here. Assuming that the transistor has $V_{BE} = 0.7$ V at $I_E = 1$ mA, what value of R_C would establish a resistor current of 1 mA? Now, to what value of β does a meter reading of full scale correspond? What is β if the meter reading is 1/5 of full scale? 1/10 of full scale?

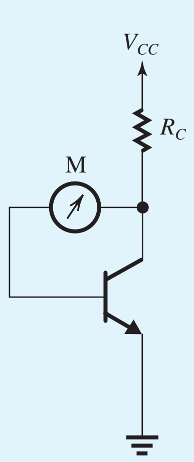


Figure P6.30

6.31 Repeat Exercise 6.13 for the situation in which the power supplies are reduced to ±2.5 V.

D 6.32 Design the circuit in Fig. P6.32 to establish a current of 0.5 mA in the emitter and a voltage of −0.5 V at the collector. The transistor $v_{EB} = 0.64$ V at $I_E = 0.1$ mA, and $\beta = 100$. To what value can R_C be increased while the collector current remains unchanged?

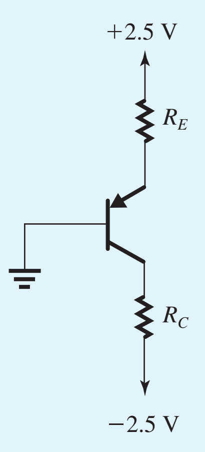


Figure P6.32

D 6.33 Examination of the table of standard values for resistors with 5% tolerance in Appendix J reveals that the closest values to those found in the design of Example 6.2 are 5.1 kΩ and 6.8 kΩ. For these values, use approximate calculations (e.g., $V_{BE} \simeq 0.7$ V and $\alpha \simeq 1$) to determine the values of collector current and collector voltage that are likely to result.

D 6.34 Design the circuit in Fig. P6.34 to establish $I_C = 0.2$ mA and $V_C = 0.5$ V. The transistor exhibits v_{BE} of 0.8 V at $i_C = 1$ mA, and $\beta = 100$.

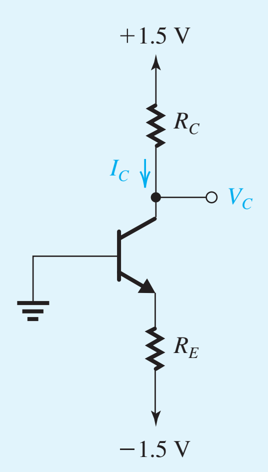


Figure P6.34

6.35 For each of the circuits shown in Fig. P6.35, find the emitter, base, and collector voltages and currents. Use $\beta = 50$, but assume $|V_{BE}| = 0.8$ V independent of current level.

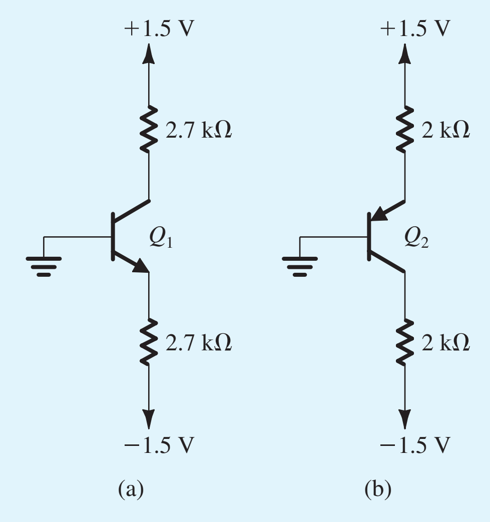


Figure P6.35

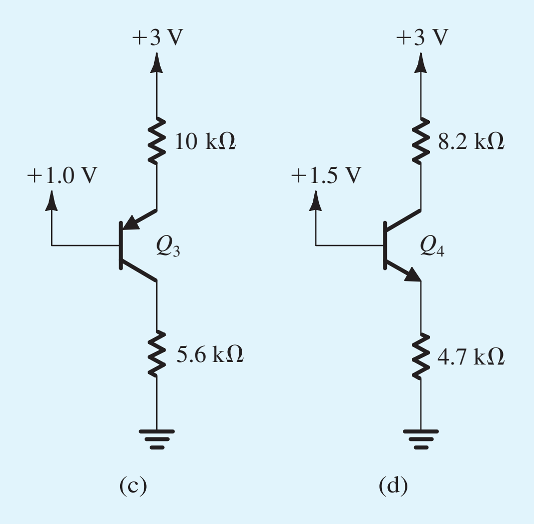


Figure P6.35 *continued*

6.36 The current I_{CBO} of a small transistor is measured to be 10 nA at 25°C. If the temperature of the device is raised to 125°C, what do you expect I_{CBO} to become?

6.37 Augment the model of the *npn* BJT shown in Fig. 6.19(a) by a current source representing I_{CBO}. Assume that r_o is very large and thus can be neglected. In terms of this addition, what do the terminal currents i_B, i_C, and i_E become? If the base lead is open-circuited while the emitter is connected to ground, and the collector is connected to a positive supply, find the emitter and collector currents.

6.38 A BJT whose emitter current is fixed at 1 mA has a base–emitter voltage of 0.70 V at 25°C. What base–emitter voltage would you expect at 0°C? At 100°C?

6.39 A particular *pnp* transistor operating at an emitter current of 0.5 mA at 20°C has an emitter–base voltage of 692 mV.

(a) What does v_{EB} become if the junction temperature rises to 50°C?
(b) If the transistor is operated at a fixed emitter–base voltage of 700 mV, what emitter current flows at 20°C? At 50°C?

6.40 Consider a transistor for which the base–emitter voltage drop is 0.7 V at 10 mA. What current flows for $v_{BE} = 0.5$ V? Evaluate the ratio of the slopes of the i_C–v_{BE} curve at $v_{BE} =$ 700 mV and at $v_{BE} = 500$ mV. The large ratio confirms the point that the BJT has an "apparent threshold" at $v_{BE} \simeq 0.5$ V.

6.41 Use Eq. (6.18) to plot i_C versus v_{CE} for an *npn* transistor having $I_S = 10^{-15}$ A and $V_A = 100$ V. Provide curves for $v_{BE} =$ 0.65, 0.70, 0.72, 0.73, and 0.74 volts. Show the characteristics for v_{CE} up to 15 V.

***6.42** In the circuit shown in Fig. P6.42, current source I is 1.1 mA, and at 25°C $v_{BE} = 680$ mV at $i_E = 1$ mA. At 25°C with $\beta = 100$, what currents flow in R_1 and R_2? What voltage would you expect at node E? Noting that the temperature coefficient of v_{BE} for I_E constant is −2 mV/°C, what is the TC of v_E? For an ambient temperature of 75°C, what voltage would you expect at node E? Clearly state any simplifying assumptions you make.

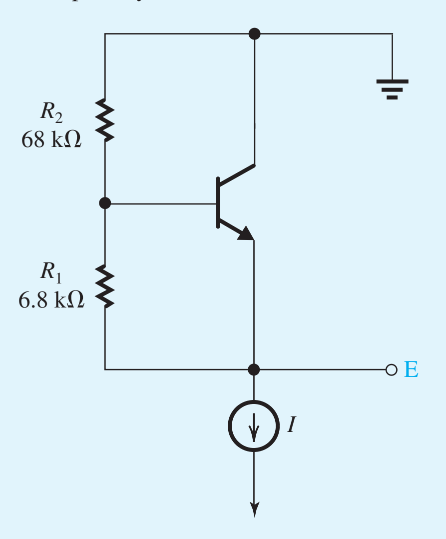


Figure P6.42

6.43 For a particular *npn* transistor operating at a v_{BE} of 680 mV and $I_C = 1$ mA, the i_C–v_{CE} characteristic has a slope of 0.8×10^{-5} ℧. To what value of output resistance does this correspond? What is the value of the Early voltage for this transistor? For operation at 10 mA, what would the output resistance become?

6.44 For a BJT having an Early voltage of 50 V, what is its output resistance at 1 mA? At 100 μA?

6.45 Measurements of the i_C–v_{CE} characteristic of a small-signal transistor operating at $v_{BE} = 710$ mV show that $i_C = 1.1$ mA at $v_{CE} = 5$ V and that $i_C = 1.3$ mA at $v_{CE} = 15$ V. What is the corresponding value of i_C near saturation? At what value of v_{CE} is $i_C = 1.2$ mA? What is the value of the Early voltage for this transistor? What is the output resistance that corresponds to operation at $v_{BE} = 710$ mV?

6.46 Give the *pnp* equivalent circuit models that correspond to those shown in Fig. 6.19 for the *npn* case.

6.47 A BJT operating at $i_B = 10\,\mu\text{A}$ and $i_C = 1.0\,\text{mA}$ undergoes a reduction in base current of 1.0 μA. It is found that when v_{CE} is held constant, the corresponding reduction in collector current is 0.08 mA. What are the values of β and the incremental β or β_{ac} that apply? If the base current is increased from 10 μA to 12 μA and v_{CE} is increased from 8 V to 10 V, what collector current results? Assume $V_A = 100$ V.

6.48 For the circuit in Fig. P6.48 let $V_{CC} = 10$ V, $R_C = 1\,\text{k}\Omega$, and $R_B = 10\,\text{k}\Omega$. The BJT has $\beta = 50$. Find the value of V_{BB} that results in the transistor operating

(a) in the active mode with $V_C = 2$ V;
(b) at the edge of saturation;
(c) deep in saturation with $\beta_{\text{forced}} = 10$.

Assume $V_{BE} \simeq 0.7$ V.

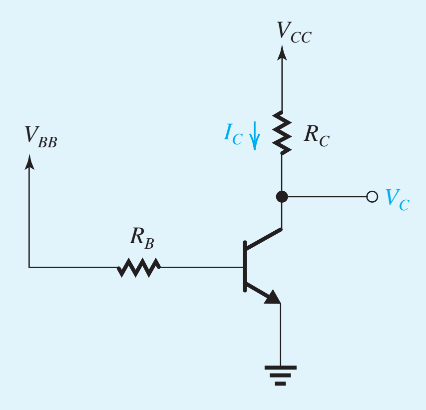


Figure P6.48

SIM **D** ***6.49** Consider the circuit of Fig. P6.48 for the case $V_{BB} = V_{CC}$. If the BJT is saturated, use the equivalent circuit of Fig. 6.21 to derive an expression for β_{forced} in terms of V_{CC} and (R_B/R_C). Also derive an expression for the total power dissipated in the circuit. For $V_{CC} = 5$ V, design the circuit to obtain operation at a forced β as close to 10 as possible while limiting the power dissipation to no larger than 20 mW. Use 1% resistors (see Appendix J).

6.50 The *pnp* transistor in the circuit in Fig. P6.50 has $\beta = 50$. Show that the BJT is operating in the saturation mode and find β_{forced} and V_C. To what value should R_B be increased in order for the transistor to operate at the edge of saturation?

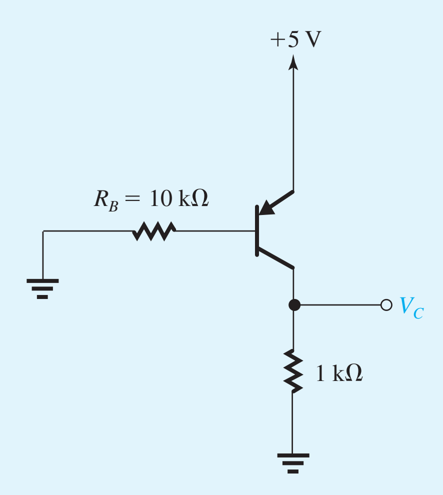


Figure P6.50

Section 6.3: BJT Circuits at DC

6.51 The transistor in the circuit of Fig. P6.51 has a very high β. Find V_E and V_C for V_B (a) +2.0 V, (b) +1.7 V, and (c) 0 V.

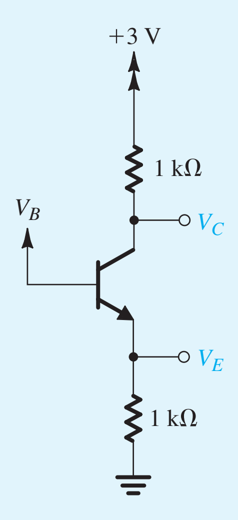


Figure P6.51

6.52 The transistor in the circuit of Fig. P6.51 has a very high β. Find the highest value of V_B for which the transistor still operates in the active mode. Also, find the value of V_B for which the transistor operates in saturation with a forced β of 2.

6.53 Consider the operation of the circuit shown in Fig. P6.53 for V_B at –1 V, 0 V, and +1 V. Assume that β is very high. What values of V_E and V_C result? At what value of V_B does the emitter current reduce to one-tenth of

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

its value for $V_B = 0$ V? For what value of V_B is the transistor just at the edge of conduction? ($v_{BE} = 0.5$ V) What values of V_E and V_C correspond? For what value of V_B does the transistor reach the edge of saturation? What values of V_C and V_E correspond? Find the value of V_B for which the transistor operates in saturation with a forced β of 2.

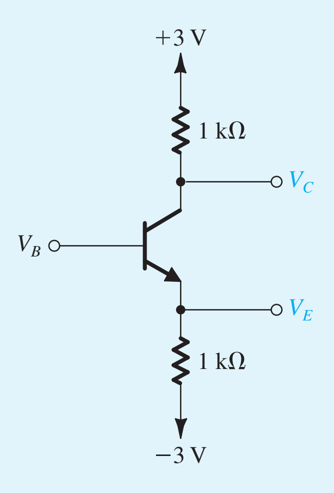


Figure P6.53

6.54 For the transistor shown in Fig. P6.54, assume $\alpha \simeq 1$ and $v_{BE} = 0.5$ V at the edge of conduction. What are the values of V_E and V_C for $V_B = 0$ V? For what value of V_B does the transistor cut off? Saturate? In each case, what values of V_E and V_C result?

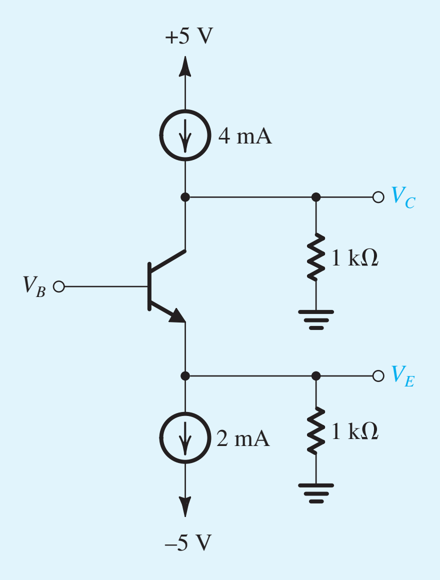


Figure P6.54

D 6.55 Consider the circuit in Fig. P6.51 with the base voltage V_B obtained using a voltage divider across the 3-V supply. Assuming the transistor β to be very large (i.e., ignoring the base current), design the voltage divider to obtain $V_B = 1.2$ V. Design for a 0.1-mA current in the voltage divider. Now, if the BJT $\beta = 100$, analyze the circuit to determine the collector current and the collector voltage.

6.56 A single measurement indicates the emitter voltage of the transistor in the circuit of Fig. P5.56 to be 1.0 V. Under the assumption that $|V_{BE}| = 0.7$ V, what are V_B, I_B, I_E, I_C, V_C, β, and α ? (*Note:* Isn't it surprising what a little measurement can lead to?)

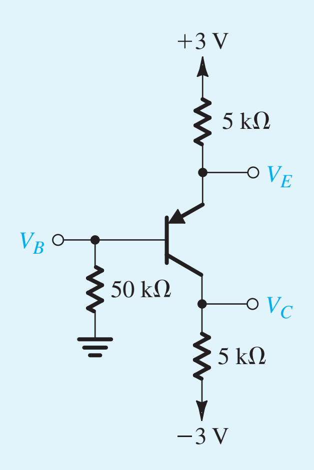


Figure P6.56

D 6.57 Design a circuit using a *pnp* transistor for which $\alpha \simeq 1$ using two resistors connected appropriately to ±3 V so that $I_E = 0.5$ mA and $V_{BC} = 1$ V. What exact values of R_E and R_C would be needed? Now, consult a table of standard 5% resistor values (e.g., that provided in Appendix J) to select suitable practical values. What values of resistors have you chosen? What are the values of I_E and V_{BC} that result?

6.58 In the circuit shown in Fig. P6.58, the transistor has $\beta = 40$. Find the values of V_B, V_E, and V_C. If R_B is raised to 100 kΩ, what voltages result? With $R_B = 100$ kΩ, what value of β would return the voltages to the values first calculated?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

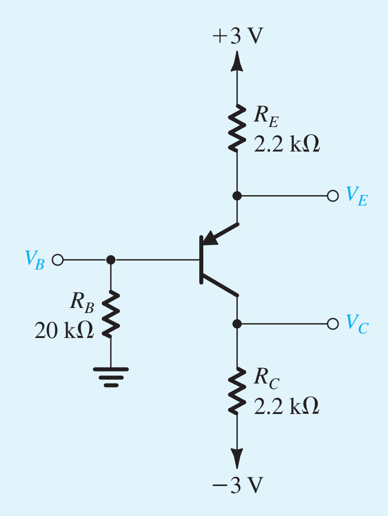


Figure P6.58

6.59 In the circuit shown in Fig. P6.58, the transistor has $\beta = 50$. Find the values of V_B, V_E, and V_C, and verify that the transistor is operating in the active mode. What is the largest value that R_C can have while the transistor remains in the active mode?

SIM **6.60** For the circuit in Fig. P6.60, find V_B, V_E, and V_C for $R_B = 100$ kΩ, 10 kΩ, and 1 kΩ. Let $\beta = 100$.

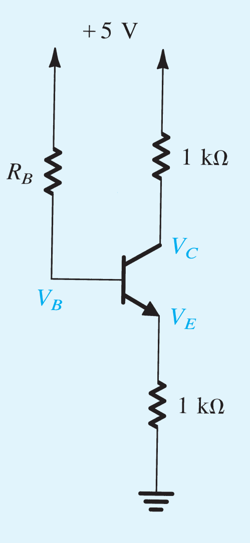
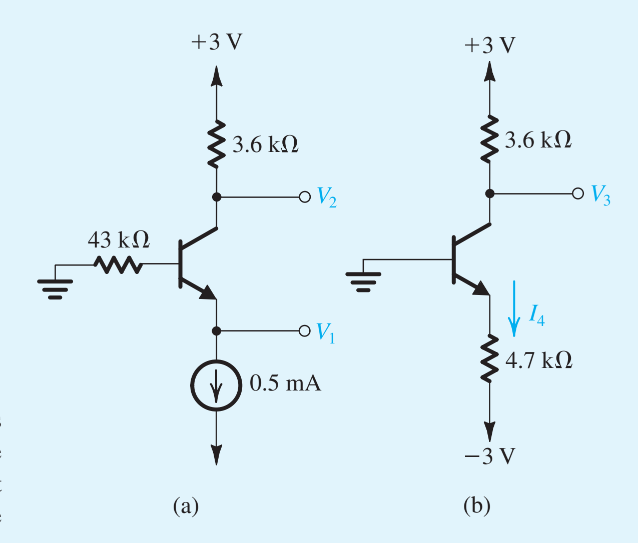


Figure P6.60

6.61 For the circuits in Fig. P6.61, find values for the labeled node voltages and branch currents. Assume β to be very high.



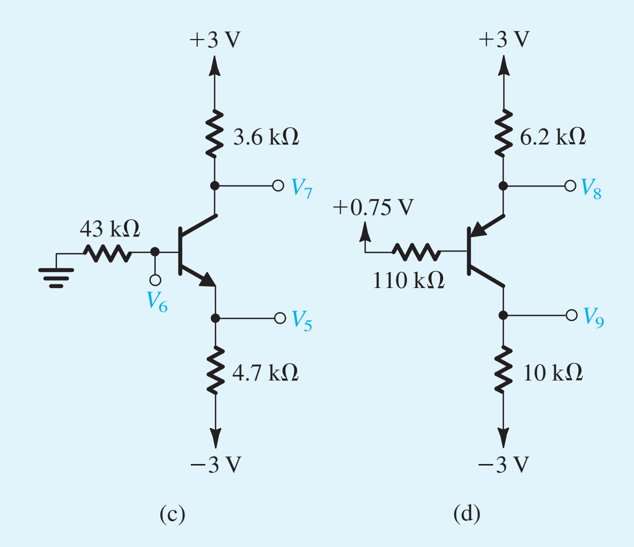


Figure P6.61

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

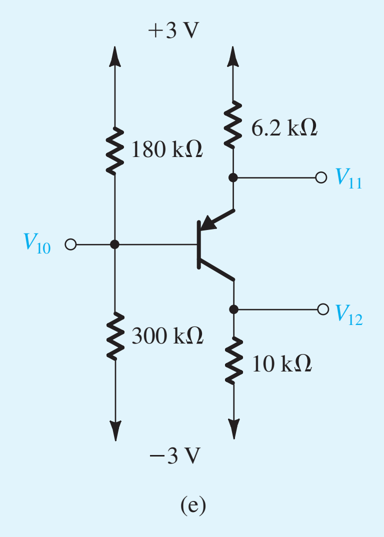


Figure P6.61 *continued*

***6.62** Repeat the analysis of the circuits in Problem 6.61 using $\beta = 100$. Find all the labeled node voltages and branch currents.

D **6.63 It is required to design the circuit in Fig. P6.63 so that a current of 1 mA is established in the emitter and a voltage of −1 V appears at the collector. The transistor type used has a nominal β of 100. However, the β value can be as low as 50 and as high as 150. Your design should ensure that the specified emitter current is obtained when $\beta = 100$ and that at the extreme values of β the emitter current does not change by more than 10% of its nominal value. Also, design for as large a value for R_B as possible. Give the values of R_B, R_E, and R_C to the nearest kilohm. What is the expected range of collector current and collector voltage corresponding to the full range of β values?

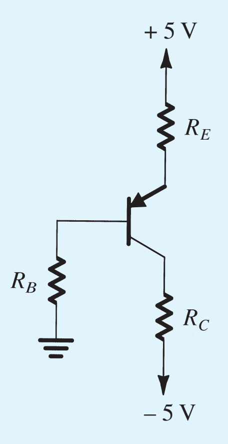


Figure P6.63

D 6.64 The *pnp* transistor in the circuit of Fig. P6.64 has $\beta = 50$. Find the value for R_C to obtain $V_C = +2$ V. What happens if the transistor is replaced with another having $\beta = 100$? Give the value of V_C in the latter case.

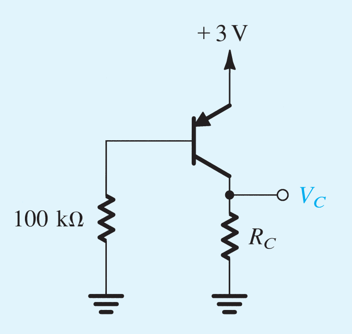


Figure P6.64

*****6.65** Consider the circuit shown in Fig. P6.65. It resembles that in Fig. 6.30 but includes other features. First, note diodes D_1 and D_2 are included to make design (and analysis) easier and to provide temperature compensation for the emitter–base voltages of Q_1 and Q_2. Second, note resistor R, whose purpose is to provide negative feedback (more on this later in the book!). Using $|V_{BE}|$ and $V_D = 0.7$ V independent of current, and $\beta = \infty$, find the voltages V_{B1}, V_{E1}, V_{C1}, V_{B2}, V_{E2}, and V_{C2}, initially with R open-circuited and then with R connected. Repeat for $\beta = 100$, with R open-circuited initially, then connected.

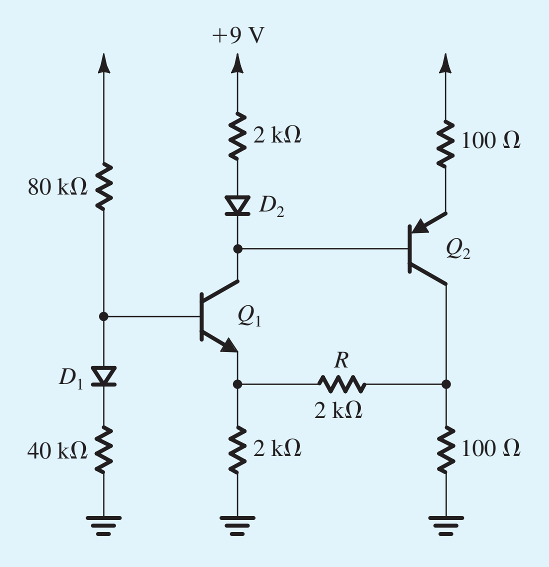


Figure P6.65

***6.66** For the circuit shown in Fig. P6.66, find the labeled node voltages for:

(a) $\beta = \infty$
(b) $\beta = 100$

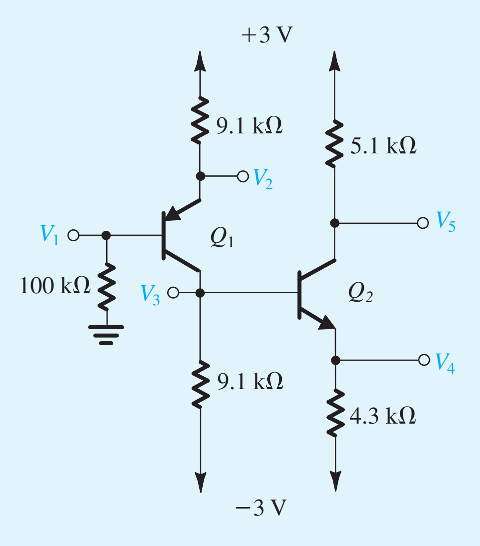


Figure P6.66

D *6.67 Using $\beta = \infty$, design the circuit shown in Fig. P6.67 so that the emitter currents of Q_1, Q_2, and Q_3 are 0.5 mA, 0.5 mA, and 1 mA, respectively, and $V_3 = 0$, $V_5 = -2$ V, and $V_7 = 1$ V. For each resistor, select the nearest standard value utilizing the table of standard values for 5% resistors in Appendix J. Now, for $\beta = 100$, find the values of V_3, V_4, V_5, V_6, and V_7.

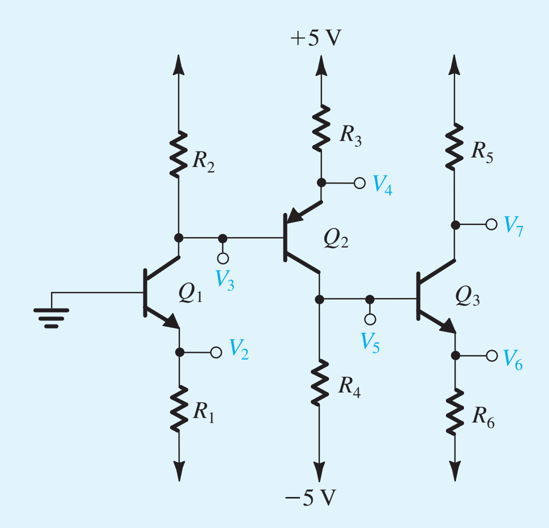


Figure P6.67

***6.68** For the circuit in Fig. P6.68, find V_B and V_E for v_I = 0 V, +2 V, –2.5 V, and –5 V. The BJTs have $\beta = 50$.

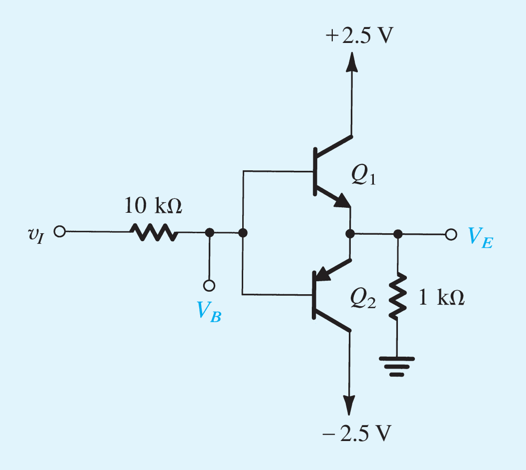


Figure P6.68

****6.69** All the transistors in the circuits of Fig. P6.69 are specified to have a minimum β of 50. Find approximate values for the collector voltages and calculate forced β for each of the transistors. (*Hint:* Initially, assume all transistors are operating in saturation, and verify the assumption.)

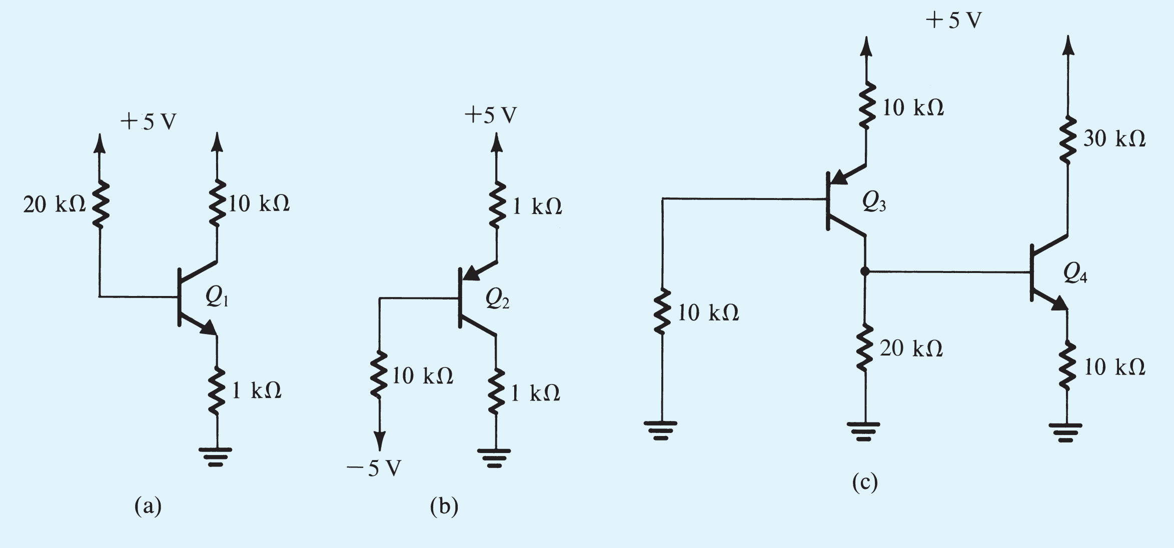

+5 V
20 kΩ
10 kΩ
Q_1
1 kΩ
(a)
+5 V
1 kΩ
Q_2
10 kΩ
1 kΩ
−5 V
(b)
+5 V
10 kΩ
Q_3
30 kΩ
Q_4
10 kΩ
20 kΩ
10 kΩ
(c)

Figure P6.69

CHAPTER 7

Transistor Amplifiers

IN THIS CHAPTER YOU WILL LEARN

1. How the transistor (a MOSFET or a BJT) can be used to make an amplifier.
2. How to obtain linear amplification from the fundamentally nonlinear MOS and bipolar transistor.
3. How to model the linear operation of a transistor around a bias point by an equivalent circuit that can be used in the analysis and design of transistor amplifiers.
4. The three basic ways to connect a MOSFET or a BJT to construct amplifiers with different properties.
5. Practical circuits for MOS and bipolar transistor amplifiers that can be constructed using discrete components.

Introduction

Having studied the two major transistor types, the MOSFET (Chapter 5) and the BJT (Chapter 6), we now begin the study of their application. There are two distinctly different kinds of transistor application: as a switch, in the design of digital circuits (Chapters 14–16) and as a controlled source, in the design of amplifiers for analog circuits. This chapter and the subsequent six focus on the latter application, namely, the use of the transistor in the design of a variety of amplifier types.

Since the basic principles that underlie the use of the MOSFET and the BJT in amplifier design are the same, the two devices are studied together in this chapter. Besides providing some economy in presentation, this unified study enables us to make important comparisons between MOS and bipolar amplifiers.

The bulk of this chapter is concerned with the fundamental principles and concepts that are the basis for the application of transistors in amplifier design: We study in detail the models that are used to represent both transistor types in the analysis and design of small-signal linear amplifiers. We also study the three basic configurations in which each of the two transistor types can be connected to realize an amplifier.

The chapter concludes with examples of discrete-circuit amplifiers. These are circuits that can be assembled using discrete transistors, resistors, and capacitors on printed-circuit boards (PCBs). They predominantly use BJTs, and their design differs in significant ways from the design of integrated-circuit (IC) amplifiers. The latter predominantly use MOSFETs, and their study begins in Chapter 8. However, the fundamental principles and concepts introduced in this chapter apply equally well to both discrete and integrated amplifiers.

7.1 Basic Principles

7.1.1 The Basis for Amplifier Operation

The basis for the application of the transistor (a MOSFET or a BJT) in amplifier design is that when the device is operated in the active region, a voltage-controlled current source is realized. Specifically, when a MOSFET is operated in the saturation or pinch-off region, also referred to in this chapter as the active region, the voltage between gate and source, v_{GS}, controls the drain current i_D according to the square-law relationship which, for an NMOS transistor, is expressed as

$$i_D = \frac{1}{2}k_n(v_{GS} - V_{tn})^2 \tag{7.1}$$

We note that in this first-order model of MOSFET operation, the drain current i_D does not depend on the drain voltage v_{DS} because the channel is pinched off at the drain end,[1] thus "isolating" the drain.

Similarly, when a BJT is operated in the active region, the base-emitter voltage v_{BE} controls the collector current i_C according to the exponential relationship which, for an *npn* transistor, is expressed as

$$i_C = I_S e^{v_{BE}/V_T} \tag{7.2}$$

Here, this first-order model of BJT operation indicates that the collector current i_C does not depend on the collector voltage v_{CE} because the collector–base junction is reverse biased, thus "isolating" the collector.

Figure 7.1 shows an NMOS transistor and an *npn* transistor operating in the active mode. Observe that for the NMOS transistor, the pinch-off condition is ensured by keeping $v_{DS} \geq v_{OV}$. Since the overdrive voltage $v_{OV} = v_{GS} - V_{tn}$, this condition implies that $v_{GD} \leq V_{tn}$, which indeed ensures channel pinch-off at the drain end.

Similarly, for the *npn* transistor in Fig. 7.1(b), the CBJ reverse-bias condition is ensured by keeping $v_{CE} \geq 0.3$ V. Since v_{BE} is usually in the vicinity of 0.7 V, v_{BC} is thus kept

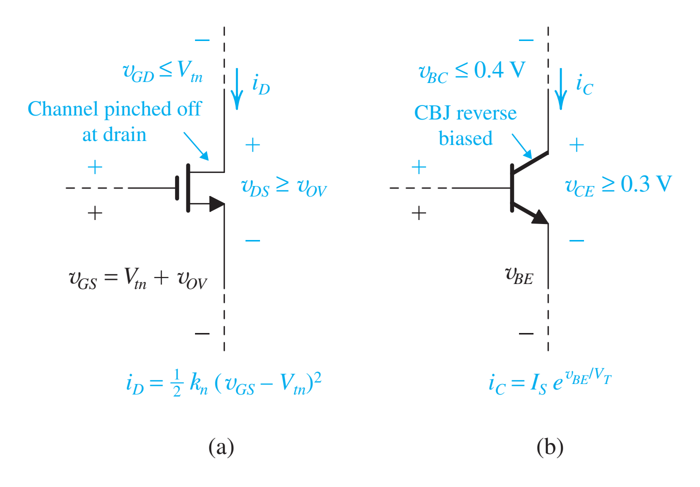


Figure 7.1 Operating (**a**) an NMOS transistor and (**b**) an *npn* transistor in the active mode. Note that $v_{GS} = V_{tn} + v_{OV}$ and $v_{DS} \geq v_{OV}$; thus $v_{GD} \leq V_{tn}$, which ensures channel pinch-off at the drain end. Similarly, $v_{BE} \simeq 0.7$ V, and $v_{CE} \geq 0.3$ V results in $v_{BC} \leq 0.4$ V, which is sufficient to keep the CBJ from conducting.

[1]To focus on essentials, we shall neglect the Early effect until a later point.

smaller than 0.4 V, which is sufficient to prevent this relatively large-area junction from conducting.

Although we used NMOS and *npn* transistors to illustrate the conditions for active-mode operation, similar conditions apply for PMOS and *pnp* transistors, as studied in Chapters 5 and 6, respectively.

Finally, we note that the control relationships in Eqs. (7.1) and (7.2) are nonlinear. Nevertheless, we shall shortly devise a technique for obtaining almost-linear amplification from these fundamentally nonlinear devices.

7.1.2 Obtaining a Voltage Amplifier

From the above we see that the transistor is basically a transconductance amplifier: that is, an amplifier whose input signal is a voltage and whose output signal is a current. More commonly, however, one is interested in voltage amplifiers. A simple way to convert a transconductance amplifier to a voltage amplifier is to pass the output current through a resistor and take the voltage across the resistor as the output. Doing this for a MOSFET results in the simple amplifier circuit shown in Fig. 7.2(a). Here v_{GS} is the input voltage, R_D (known as a **load resistance**) converts the drain current i_D to a voltage ($i_D R_D$), and V_{DD} is the supply voltage that powers up the amplifier and, together with R_D, establishes operation in the active region, as will be shown shortly.

In the amplifier circuit of Fig. 7.2(a) the output voltage is taken between the drain and ground, rather than simply across R_D. This is done because of the need to maintain a common ground reference between the input and the output. The output voltage v_{DS} is given by

$$v_{DS} = V_{DD} - i_D R_D \tag{7.3}$$

Thus it is an inverted version (note the minus sign) of $i_D R_D$ that is shifted by the constant value of the supply voltage V_{DD}.

An exactly similar arrangement applies for the BJT amplifier, as illustrated in Fig. 7.2(c). Here the output voltage v_{CE} is given by

$$v_{CE} = V_{CC} - i_C R_C \tag{7.4}$$

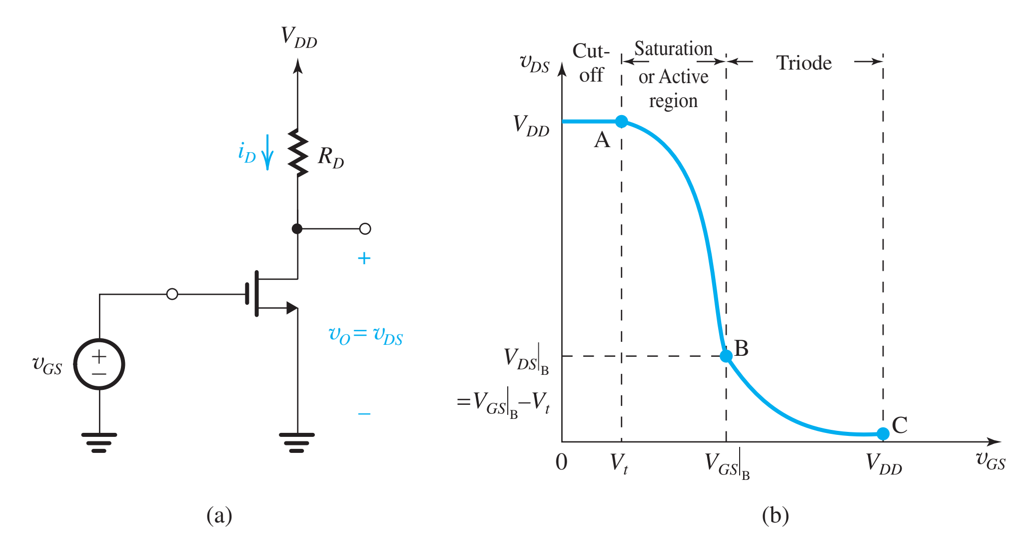


Figure 7.2 (**a**) An NMOS amplifier and (**b**) its VTC; and (**c**) an *npn* amplifier and (**d**) its VTC.

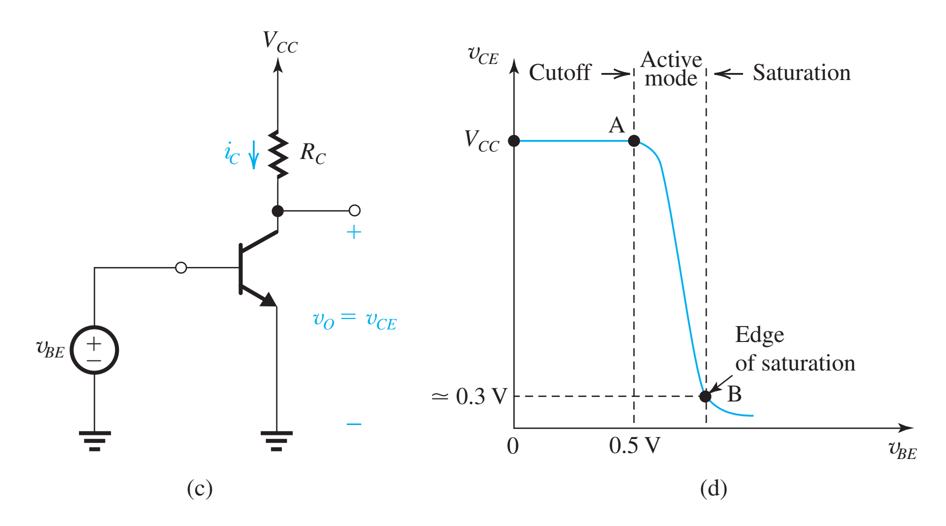


Figure 7.2 *continued*

7.1.3 The Voltage-Transfer Characteristic (VTC)

A useful tool that provides insight into the operation of an amplifier circuit is its voltage-transfer characteristic (VTC). This is simply a plot (or a clearly labeled sketch) of the output voltage versus the input voltage. For the MOS amplifier in Fig. 7.2(a), this is the plot of v_{DS} versus v_{GS} shown in Fig. 7.2(b).

Observe that for $v_{GS} < V_t$, the transistor is cut off, $i_D = 0$ and, from Eq. (7.3), $v_{DS} = V_{DD}$. As v_{GS} exceeds V_t, the transistor turns on and v_{DS} decreases. However, since initially v_{DS} is still high, the MOSFET will be operating in saturation or the active region. This continues as v_{GS} is increased until the value of v_{GS} is reached that results in v_{DS} becoming lower than v_{GS} by V_t volts [point B on the VTC in Fig. 7.2(b)]. For v_{GS} greater than that at point B, the transistor operates in the triode region and v_{DS} decreases more slowly.

The VTC in Fig. 7.2(b) indicates that the segment of greatest slope (hence potentially the largest amplifier gain) is that labeled AB, which corresponds to operation in the active region. When a MOSFET is operated as an amplifier, its operating point is confined to the segment AB at all times. An expression for the segment AB can be obtained by substituting for i_D in Eq. (7.3) by its active-region value from Eq. (7.1), thus

$$v_{DS} = V_{DD} - \frac{1}{2}k_n R_D (v_{GS} - V_t)^2 \tag{7.5}$$

This is obviously a nonlinear relationship. Nevertheless, linear (or almost-linear) amplification can be obtained by using the technique of biasing the MOSFET. Before considering biasing, however, it is useful to determine the coordinates of point B, which is at the boundary between the saturation and the triode regions of operation. These can be obtained by substituting in Eq. (7.5), $v_{GS} = V_{GS}|_B$ and $v_{DS} = V_{DS}|_B = V_{GS}|_B - V_t$. The result is

$$V_{GS}|_B = V_t + \frac{\sqrt{2k_n R_D V_{DD} + 1} - 1}{k_n R_D} \tag{7.6}$$

Point B can be alternatively characterized by the overdrive voltage

$$V_{OV}\big|_{B} \equiv V_{GS}\big|_{B} - V_t = \frac{\sqrt{2k_n R_D V_{DD} + 1} - 1}{k_n R_D} \tag{7.7}$$

and

$$V_{DS}\big|_{B} = V_{OV}\big|_{B} \tag{7.8}$$

EXERCISE

7.1 Consider the amplifier of Fig. 7.2(a) with $V_{DD} = 1.8$ V, $R_D = 17.5$ kΩ, and with a MOSFET specified to have $V_t = 0.4$ V, $k_n = 4$ mA/V^2, and $\lambda = 0$. Determine the coordinates of the end points of the active-region segment of the VTC. Also, determine $V_{DS}\big|_C$ assuming $V_{GS}\big|_C = V_{DD}$.
Ans. A: 0.4 V, 1.8 V; B: 0.613 V, 0.213 V; $V_{DS}\big|_C = 18$ mV

An exactly similar development applies to the BJT case. This is illustrated in Fig. 7.2(c) and (d). In this case, over the active-region or amplifier segment AB, the output voltage v_{CE} is related to the input voltage v_{BE} by

$$v_{CE} = V_{CC} - R_C I_S e^{v_{BE}/V_T} \tag{7.9}$$

Here also, the input–output relationship is nonlinear. Nevertheless, linear (or almost-linear) amplification can be obtained by using the biasing technique discussed next.

7.1.4 Obtaining Linear Amplification by Biasing the Transistor

Biasing enables us to obtain almost-linear amplification from the MOSFET and the BJT. The technique is illustrated for the MOSFET case in Fig. 7.3(a). A dc voltage V_{GS} is selected to obtain operation at a point Q on the segment AB of the VTC. How to select an appropriate location for the bias point Q will be discussed shortly. For the time being, observe that the coordinates of Q are the dc voltages V_{GS} and V_{DS}, which are related by

$$V_{DS} = V_{DD} - \frac{1}{2}k_n R_D (V_{GS} - V_t)^2 \tag{7.10}$$

Point Q is known as the **bias point** or the **dc operating point**. Also, since at Q no signal component is present, it is also known as the **quiescent point** (which is the origin of the symbol Q).

Next, the signal to be amplified, v_{gs}, a function of time t, is superimposed on the bias voltage V_{GS}, as shown in Fig. 7.4(a). Thus the total instantaneous value of v_{GS} becomes

$$v_{GS}(t) = V_{GS} + v_{gs}(t)$$

The resulting $v_{DS}(t)$ can be obtained by substituting for $v_{GS}(t)$ into Eq. (7.5). Graphically, we can use the VTC to obtain $v_{DS}(t)$ point by point, as illustrated in Fig. 7.4(b). Here we show

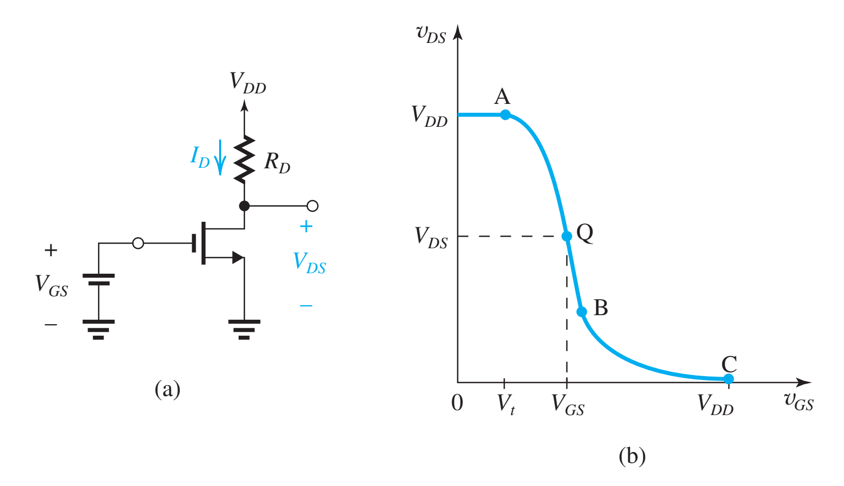


Figure 7.3 Biasing the MOSFET amplifier at a point Q located on the segment AB of the VTC.

the case of v_{gs} being a triangular wave of "small" amplitude. Specifically, the amplitude of v_{gs} is small enough to restrict the excursion of the instantaneous operating point to a short, almost-linear segment of the VTC around the bias point Q. The shorter the segment, the greater the linearity achieved, and the closer to an ideal triangular wave the signal component at the output, v_{ds}, will be. This is the essence of obtaining linear amplification from the nonlinear MOSFET.

Before leaving Fig. 7.4(b) we wish to draw the reader's attention to the consequence of increasing the amplitude of the signal v_{gs}. As the instantaneous operating point will no longer be confined to the almost-linear segment of the VTC, the output signal v_{ds} will deviate from its ideal triangular shape; that is, it will exhibit nonlinear distortion. Worse yet, if the input signal amplitude becomes sufficiently large, the instantaneous operating point may leave the segment AB altogether. If this happens at the negative peaks of v_{gs}, the transistor will cut off for a portion of the cycle and the positive peaks of v_{ds} will be "clipped off." If it occurs at the positive peaks of v_{gs}, the transistor will enter the triode region for a portion of the cycle, and the negative peaks of v_{ds} will become flattened. It follows that the selection of the location of the bias point Q can have a profound effect on the maximum allowable amplitude of v_{ds}, referred to as the *allowable signal swing at the output*. We will have more to say later on this important point.

An exactly parallel development can be applied to the BJT amplifier. In fact, all we need to do is replace the NMOS transistor in Figs. 7.3 and 7.4 with an *npn* transistor and change the voltage and current symbols to their BJT counterparts. The resulting bias point Q will be characterized by dc voltages V_{BE} and V_{CE}, which are related by

$$V_{CE} = V_{CC} - R_C I_S e^{V_{BE}/V_T} \tag{7.11}$$

and a dc current I_C,

$$I_C = I_S e^{V_{BE}/V_T} \tag{7.12}$$

Also, superimposing a small-signal v_{be} on the dc bias voltage V_{BE} results in

$$v_{BE}(t) = V_{BE} + v_{be}(t)$$

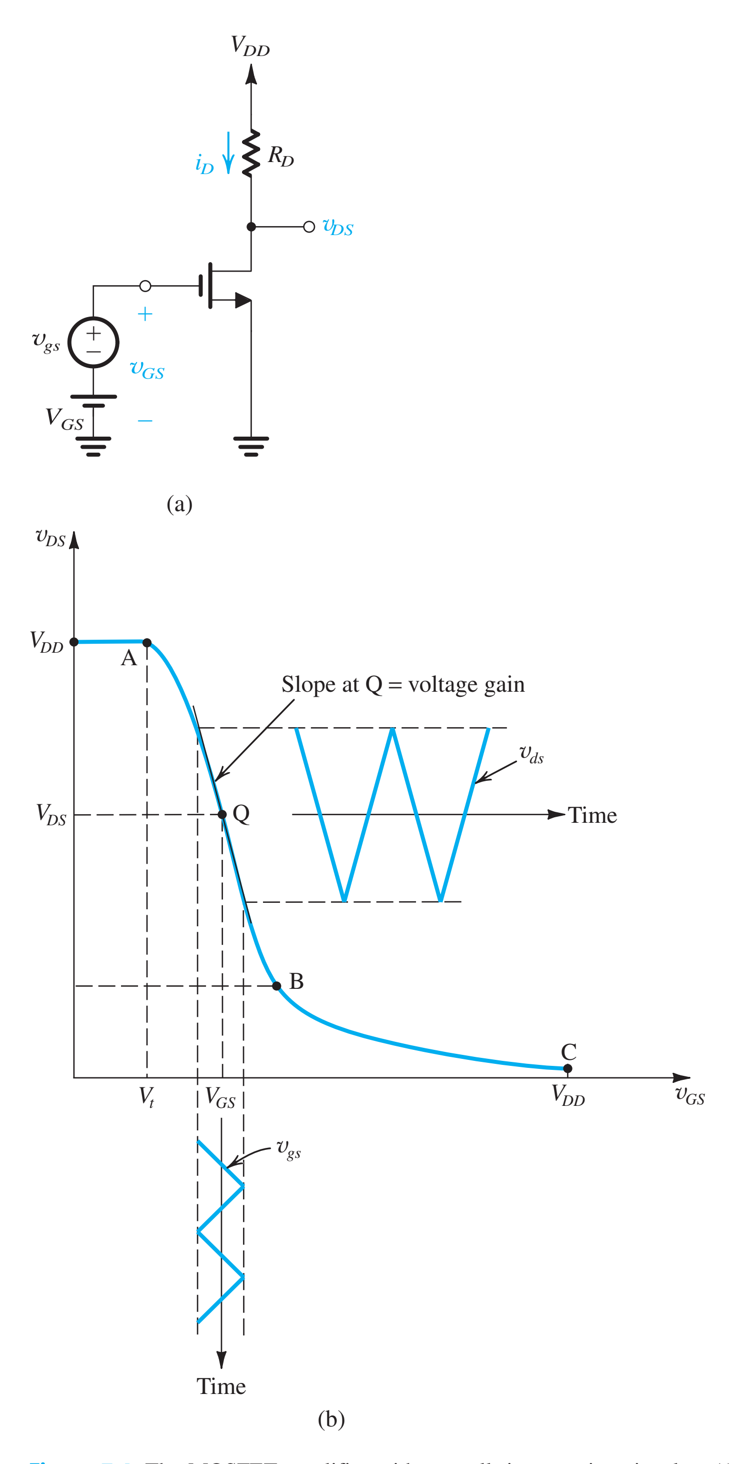


Figure 7.4 The MOSFET amplifier with a small time-varying signal $v_{gs}(t)$ superimposed on the dc bias voltage V_{GS}. The MOSFET operates on a short almost-linear segment of the VTC around the bias point Q and provides an output voltage $v_{ds} = A_v v_{gs}$.

which can be substituted into Eq. (7.9) to obtain the total instantaneous value of the output voltage $v_{CE}(t)$. Here again, almost-linear operation is obtained by keeping v_{be} small enough to restrict the excursion of the instantaneous operating point to a short, almost-linear segment of the VTC around the bias point Q. Similar comments also apply to the maximum allowable signal swing at the output.

7.1.5 The Small-Signal Voltage Gain

The MOSFET Case Consider the MOSFET amplifier in Fig. 7.4(a). If the input signal v_{gs} is kept small, the corresponding signal at the output v_{ds} will be nearly proportional to v_{gs} with the constant of proportionality being the slope of the almost-linear segment of the VTC around Q. This is the voltage gain of the amplifier, and its value can be determined by evaluating the slope of the tangent to the VTC at the bias point Q,

$$A_v = \left.\frac{dv_{DS}}{dv_{GS}}\right|_{v_{GS}=V_{GS}} \tag{7.13}$$

Utilizing Eq. (7.5) we obtain

$$A_v = -k_n(V_{GS} - V_t)R_D \tag{7.14}$$

which can be expressed in terms of the overdrive voltage at the bias point, V_{OV}, as

$$A_v = -k_n V_{OV} R_D \tag{7.15}$$

We make the following observations on this expression for the voltage gain.

1. The gain is negative, which signifies that the amplifier is inverting; that is, there is a 180° phase shift between the input and the output. This inversion is obvious in Fig. 7.4(b) and should have been anticipated from Eq. (7.5).
2. The gain is proportional to the load resistance R_D, to the transistor transconductance parameter k_n, and to the overdrive voltage V_{OV}. This all makes intuitive sense.

Another simple and insightful expression for the voltage gain A_v can be derived by recalling that the dc current in the drain at the bias point is related to V_{OV} by

$$I_D = \frac{1}{2}k_n V_{OV}^2$$

This equation can be combined with Eq. (7.15) to obtain

$$A_v = -\frac{I_D R_D}{V_{OV}/2} \tag{7.16}$$

That is, the gain is simply the ratio of the dc voltage drop across the load resistance R_D to $V_{OV}/2$. It can be expressed in the alternative form

$$A_v = -\frac{V_{DD} - V_{DS}}{V_{OV}/2} \tag{7.17}$$

Since the maximum slope of the VTC in Fig. 7.4(b) occurs at point B, the maximum gain magnitude $|A_{v\max}|$ is obtained by biasing the transistor at point B,

$$|A_{v\max}| = \frac{V_{DD} - V_{DS}|_{\mathrm{B}}}{V_{OV}|_{\mathrm{B}}/2}$$

and since $V_{DS}\big|_B = V_{OV}\big|_B$,

$$|A_{v\max}| = \frac{V_{DD} - V_{OV}\big|_B}{V_{OV}\big|_B/2} \tag{7.18}$$

where $V_{OV}\big|_B$ is given by Eq. (7.7). Of course, this result is only of theoretical importance since biasing at B would leave no room for negative signal swing at the output. Nevertheless, the result in Eq. (7.18) is valuable as it provides an upper bound on the magnitude of voltage gain achievable from this basic amplifier circuit. As an example, for a discrete-circuit amplifier operated with $V_{DD} = 5$ V and $V_{OV}\big|_B = 0.5$ V, the maximum achievable gain is 18 V/V. An integrated-circuit amplifier utilizing a modern submicron MOSFET operated with $V_{DD} = 1.3$ V and with $V_{OV}\big|_B = 0.2$ V realizes a maximum gain of 11 V/V.

Finally, note that to maximize the gain, the bias point Q should be as close to point B as possible, consistent with the required signal swing at the output. This point will be explored further in the end-of-chapter problems.

Example 7.1

Consider the amplifier circuit shown in Fig. 7.4(a). The transistor is specified to have $V_t = 0.4$ V, $k_n' = 0.4$ mA/V^2, $W/L = 10$, and $\lambda = 0$. Also, let $V_{DD} = 1.8$ V, $R_D = 17.5$ kΩ, and $V_{GS} = 0.6$ V.

(a) For $v_{gs} = 0$ (and hence $v_{ds} = 0$), find V_{OV}, I_D, V_{DS}, and A_v.
(b) What is the maximum symmetrical signal swing allowed at the drain? Hence, find the maximum allowable amplitude of a sinusoidal v_{gs}.

Solution

(a) With $V_{GS} = 0.6$ V, $V_{OV} = 0.6 - 0.4 = 0.2$ V. Thus,

$$I_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)V_{OV}^2$$
$$= \frac{1}{2} \times 0.4 \times 10 \times 0.2^2 = 0.08 \text{ mA}$$
$$V_{DS} = V_{DD} - R_D I_D$$
$$= 1.8 - 17.5 \times 0.08 = 0.4 \text{ V}$$

Since V_{DS} is greater than V_{OV}, the transistor is indeed operating in saturation. The voltage gain can be found from Eq. (7.15),

$$A_v = -k_n V_{OV} R_D$$
$$= -0.4 \times 10 \times 0.2 \times 17.5$$
$$= -14 \text{ V/V}$$

An identical result can be found using Eq. (7.17).

(b) Since $V_{OV} = 0.2$ V and $V_{DS} = 0.4$ V, we see that the maximum allowable negative signal swing at the drain is 0.2 V. In the positive direction, a swing of +0.2 V would not cause the transistor to

Example 7.1 *continued*

cut off (since the resulting v_{DS} would be still lower than V_{DD}) and thus is allowed. Thus the maximum symmetrical signal swing allowable at the drain is ±0.2 V. The corresponding amplitude of v_{gs} can be found from

$$\hat{v}_{gs} = \frac{\hat{v}_{ds}}{|A_v|} = \frac{0.2 \text{ V}}{14} = 14.2 \text{ mV}$$

Since $\hat{v}_{gs} \ll V_{OV}$, the operation will be reasonably linear (more on this in later sections).

Greater insight into the issue of allowable signal swing can be obtained by examining the signal waveforms shown in Fig. 7.5. Note that for the MOSFET to remain in saturation at the negative peak of v_{ds}, we must ensure that

$$v_{DS\text{min}} \geq v_{GS\text{max}} - V_t$$

that is,

$$0.4 - |A_v|\hat{v}_{gs} \geq 0.6 + \hat{v}_{gs} - 0.4$$

which results in

$$\hat{v}_{gs} \leq \frac{0.2}{|A_v| + 1} = 13.3 \text{ mV}$$

This result differs slightly from the one obtained earlier.

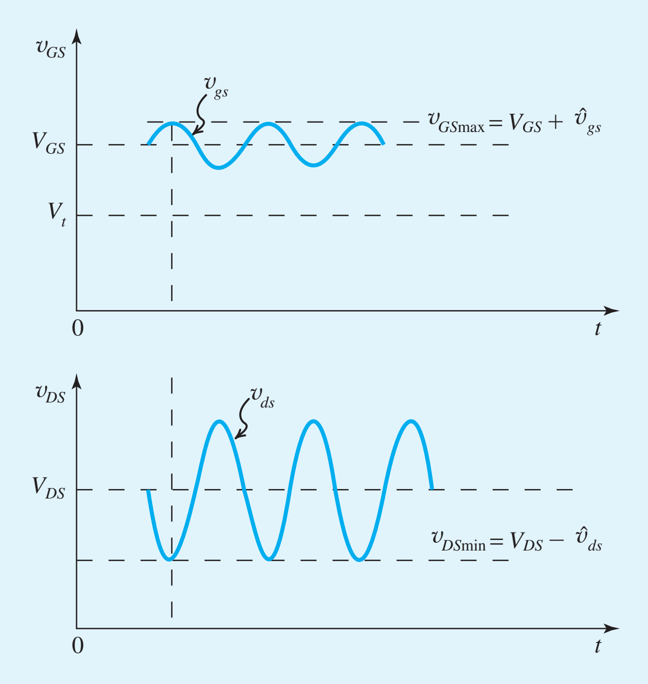


Figure 7.5 Signal waveforms at gate and drain for the amplifier in Example 7.1. Note that to ensure operation in the saturation region at all times, $v_{DS\text{min}} \geq v_{GS\text{max}} - V_t$.

EXERCISE

D7.2 For the amplifier circuit studied in Example 7.1, create two alternative designs, each providing a voltage gain of −10 by (a) changing R_D while keeping V_{OV} constant and (b) changing V_{OV} while keeping R_D constant. For each design, specify V_{GS}, I_D, R_D, and V_{DS}.
Ans. (a) 0.6 V, 0.08 mA, 12.5 kΩ, 0.8 V; (b) 0.54 V, 0.04 mA, 17.5 kΩ, 1.1 V

The BJT Case A similar development can be used to obtain the small-signal voltage gain of the BJT amplifier shown in Fig. 7.6,

$$A_v = \left.\frac{dv_{CE}}{dv_{BE}}\right|_{v_{BE}=V_{BE}} \tag{7.19}$$

Utilizing Eq. (7.9) together with Eq. (7.12), we obtain

$$A_v = -\left(\frac{I_C}{V_T}\right)R_C \tag{7.20}$$

We make the following observations on this expression for the voltage gain:

1. The gain is negative, which signifies that the amplifier is inverting; that is, there is a 180° phase shift between the input and the output. This inversion should have been anticipated from Eq. (7.9).
2. The gain is proportional to the collector bias current I_C and to the load resistance R_C.

Additional insight into the voltage gain A_v can be obtained by expressing Eq. (7.20) as

$$A_v = -\frac{I_C R_C}{V_T} \tag{7.21}$$

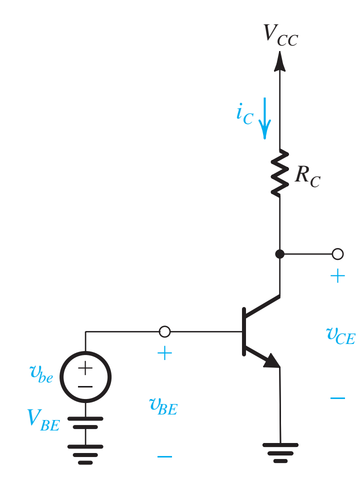


Figure 7.6 BJT amplifier biased at a point Q, with a small voltage signal v_{be} superimposed on the dc bias voltage V_{BE}. The resulting output signal v_{ce} appears superimposed on the dc collector voltage V_{CE}. The amplitude of v_{ce} is larger than that of v_{be} by the voltage gain A_v.

That is, the gain is the ratio of the dc voltage drop across the load resistance R_C to the physical constant V_T (recall that the thermal voltage $V_T \simeq 25$ mV at room temperature). This relationship is similar in form to that for the MOSFET (Eq. 7.16) except that here the denominator is a physical constant (V_T) rather than a design parameter ($V_{OV}/2$). Usually, $V_{OV}/2$ is larger than (V_T), thus we can obtain higher voltage gain from the BJT amplifier than from the MOSFET amplifier. This should not be surprising, as the exponential i_C–v_{BE} relationship is much steeper than the square-law relationship i_D–v_{GS}.

The gain A_v in Eq. (7.21) can be expressed alternately as

$$A_v = -\frac{V_{CC} - V_{CE}}{V_T} \tag{7.22}$$

from which we see that maximum gain is achieved when V_{CE} is at its minimum value of about 0.3 V,

$$|A_{v\max}| = \frac{V_{CC} - 0.3}{V_T} \tag{7.23}$$

Here again, this is only a theoretical maximum, since biasing the BJT at the edge of saturation leaves no room for negative signal swing at the output. Equation (7.23) nevertheless provides an upper bound on the voltage gain achievable from the basic BJT amplifier. As an example, for $V_{CC} = 5$ V, the maximum gain is 188 V/V, considerably larger than in the MOSFET case. For modern low-voltage technologies, a V_{CC} of 1.3 V provides a gain of 40 V/V, again much larger than the MOSFET case. The reader should not, however, jump to the conclusion that the BJT is preferred to the MOSFET in the design of modern integrated-circuit amplifiers; in fact, the opposite is true, as we shall see in Chapter 8 and beyond.

Finally, we conclude from Eq. (7.22) that to maximize $|A_v|$ the transistor should be biased at the lowest possible V_{CE} consistent with the desired value of negative signal swing at the output.

Example 7.2

Consider an amplifier circuit using a BJT having $I_S = 10^{-15}$ A, a collector resistance $R_C = 6.8$ kΩ, and a power supply $V_{CC} = 10$ V.

(a) Determine the value of the bias voltage V_{BE} required to operate the transistor at $V_{CE} = 3.2$ V. What is the corresponding value of I_C?

(b) Find the voltage gain A_v at this bias point. If an input sine-wave signal of 5-mV peak amplitude is superimposed on V_{BE}, find the amplitude of the output sine-wave signal (assume linear operation).

(c) Find the positive increment in v_{BE} (above V_{BE}) that drives the transistor to the edge of saturation, where $v_{CE} = 0.3$ V.

(d) Find the negative increment in v_{BE} that drives the transistor to within 1% of cutoff (i.e., to $v_{CE} = 0.99V_{CC}$).

Solution

(a)

$$I_C = \frac{V_{CC} - V_{CE}}{R_C}$$
$$= \frac{10 - 3.2}{6.8} = 1 \text{ mA}$$

The value of V_{BE} can be determined from

$$1 \times 10^{-3} = 10^{-15} e^{V_{BE}/V_T}$$

which results in

$$V_{BE} = 690.8 \text{ mV}$$

(b)

$$A_v = -\frac{V_{CC} - V_{CE}}{V_T}$$
$$= \frac{10 - 3.2}{0.025} = -272 \text{ V/V}$$
$$\hat{v}_{ce} = 272 \times 0.005 = 1.36 \text{ V}$$

(c) For $v_{CE} = 0.3$ V,

$$i_C = \frac{10 - 0.3}{6.8} = 1.617 \text{ mA}$$

To increase i_C from 1 mA to 1.617 mA, v_{BE} must be increased by

$$\triangle v_{BE} = V_T \ln\left(\frac{1.617}{1}\right)$$
$$= 12 \text{ mV}$$

(d) For $v_{CE} = 0.99V_{CC} = 9.9$ V,

$$i_C = \frac{10 - 9.9}{6.8} = 0.0147 \text{ mA}$$

To decrease i_C from 1 mA to 0.0147 mA, v_{BE} must change by

$$\triangle v_{BE} = V_T \ln\left(\frac{0.0147}{1}\right)$$
$$= -105.5 \text{ mV}$$

EXERCISE

7.3 For the situation described in Example 7.2, while keeping I_C unchanged at 1 mA, find the value of R_C that will result in a voltage gain of −320 V/V. What is the largest negative signal swing allowed at the output (assume that v_{CE} is not to decrease below 0.3 V)? What (approximately) is the corresponding input signal amplitude? (Assume linear operation.)
Ans. 8 kΩ; 1.7 V; 5.3 mV

7.1.6 Determining the VTC by Graphical Analysis

Figure 7.7 shows a graphical method for determining the VTC of the amplifier of Fig. 7.4(a). Although graphical analysis of transistor circuits is rarely employed in practice, it is useful to us at this stage for gaining greater insight into circuit operation, especially in answering the question of where to locate the bias point Q.

The graphical analysis is based on the observation that for each value of v_{GS}, the circuit will be operating at the point of intersection of the i_D−v_{DS} graph corresponding to the particular

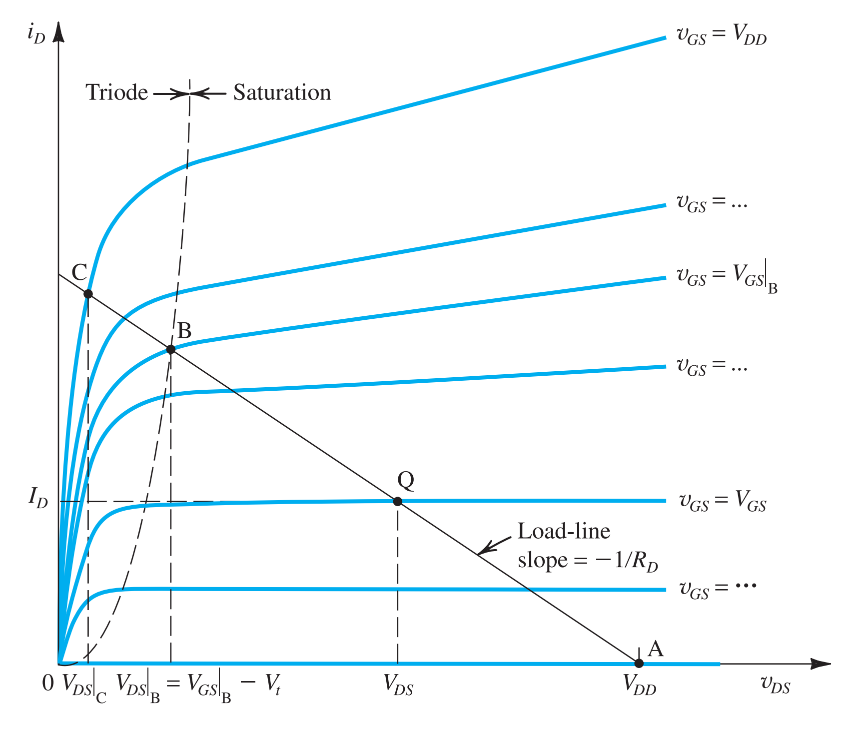


Figure 7.7 Graphical construction to determine the voltage-transfer characteristic of the amplifier in Fig. 7.4(a).

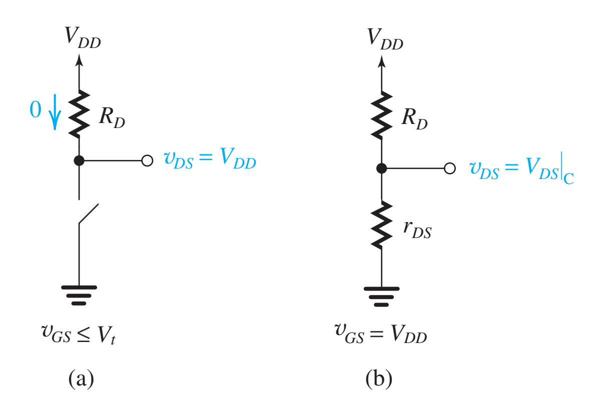


Figure 7.8 Operation of the MOSFET in Fig. 7.4(a) as a switch: **(a)** open, corresponding to point A in Fig. 7.7; **(b)** closed, corresponding to point C in Fig. 7.7. The closure resistance is approximately equal to r_{DS} because V_{DS} is usually very small.

value of v_{GS} and the straight line representing Eq. (7.3), which can be rewritten in the form

$$i_D = \frac{V_{DD}}{R_D} - \frac{1}{R_D} v_{DS} \tag{7.24}$$

The straight line representing this relationship is superimposed on the i_D-v_{DS} characteristics in Fig. 7.7. It intersects the horizontal axis at $v_{DS} = V_{DD}$ and has a slope of $-1/R_D$. Since this straight line represents in effect the load resistance R_D, it is called the **load line**. The VTC is then determined point by point. Note that we have labeled four important points: point A at which $v_{GS} = V_t$, point Q at which the MOSFET can be biased for amplifier operation ($v_{GS} = V_{GS}$ and $v_{DS} = V_{DS}$), point B at which the MOSFET leaves saturation and enters the triode region, and point C, which is deep into the triode region and for which $v_{GS} = V_{DD}$. If the MOSFET is to be used as a switch, then operating points A and C are applicable: At A the transistor is off (open switch), and at C the transistor operates as a low-valued resistance r_{DS} and has a small voltage drop (closed switch). The incremental resistance at point C is also known as the **closure resistance**. The operation of the MOSFET as a switch is illustrated in Fig. 7.8. A detailed study of the application of the MOSFET as a switch is undertaken in Chapter 14, dealing with CMOS digital logic circuits.

The graphical analysis method above can be applied to determine the VTC of the BJT amplifier in Fig. 7.2(c). Here point A, Fig. 7.2(d), corresponds to the BJT just turning on ($v_{BE} \simeq 0.5$ V) and point B corresponds to the BJT leaving the active region and entering the saturation region. If the BJT is to be operated as a switch, the two modes of operation are cutoff (open switch) and saturation (closed switch). As discussed in Section 6.2, in saturation, the BJT has a small closure resistance R_{CEsat} as well as an offset voltage. More seriously, switching the BJT out of its saturation region can require a relatively long delay time to ensure the removal of the charge stored in the BJT base region. This phenomenon has made the BJT much less attractive in digital logic applications relative to the MOSFET.[2]

7.1.7 Deciding on a Location for the Bias Point Q

For the MOSFET amplifier, the bias point Q is determined by the value of V_{GS} and that of the load resistance R_D. Two important considerations in deciding on the location of Q

[2]The only exception is a nonsaturating form of BJT logic circuits known as emitter-coupled logic (ECL).

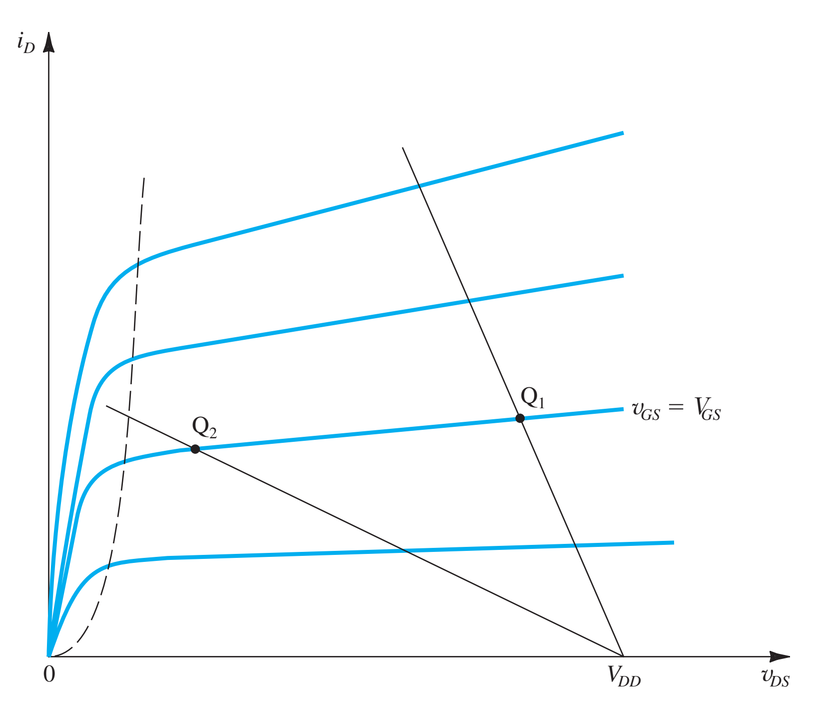


Figure 7.9 Two load lines and corresponding bias points. Bias point Q_1 does not leave sufficient room for positive signal swing at the drain (too close to V_{DD}). Bias point Q_2 is too close to the boundary of the triode region and might not allow for sufficient negative signal swing.

are the required gain and the desired signal swing at the output. To illustrate, consider the VTC shown in Fig. 7.4(b). Here the value of R_D is fixed and the only variable remaining is the value of V_{GS}. Since the slope increases as we move closer to point B, we obtain higher gain by locating Q as close to B as possible. However, the closer Q is to the boundary point B, the smaller the allowable magnitude of negative signal swing. Thus, as often happens in engineering design, we encounter a situation requiring a trade-off. The answer here is relatively simple: For a given R_D, locate Q as close to the triode region (point B) as possible to obtain high gain but sufficiently distant to allow for the required negative signal swing.

In deciding on a value for R_D, it is useful to refer to the $i_D - v_{DS}$ plane. Figure 7.9 shows two load lines resulting in two extreme bias points: Point Q_1 is too close to V_{DD}, resulting in a severe constraint on the positive signal swing of v_{ds}. Exceeding the allowable positive maximum results in the positive peaks of the signal being clipped off, since the MOSFET will turn off for the part of each cycle near the positive peak. We speak of this situation by saying that the circuit does not have sufficient "headroom." Similarly, point Q_2 is too close to the boundary of the triode region, thus severely limiting the allowable negative signal swing of v_{ds}. Exceeding this limit would result in the transistor entering the triode region for part of each cycle near the negative peaks, resulting in a distorted output signal. In this situation we say that the circuit does not have sufficient "legroom." We will have more to say on bias design in Section 7.4.

Finally, we note that exactly similar considerations apply to the case of the BJT amplifier.

7.2 Small-Signal Operation and Models

In our study of the operation of the MOSFET and BJT amplifiers in Section 7.1 we learned that linear amplification can be obtained by biasing the transistor to operate in the active region and by keeping the input signal small. In this section, we explore the small-signal operation in greater detail.

7.2.1 The MOSFET Case

Consider the conceptual amplifier circuit shown in Fig. 7.10. Here the MOS transistor is biased by applying a dc voltage[3] V_{GS}, and the input signal to be amplified, v_{gs}, is superimposed on the dc bias voltage V_{GS}. The output voltage is taken at the drain.

The DC Bias Point The dc bias current I_D can be found by setting the signal v_{gs} to zero; thus,

$$I_D = \frac{1}{2}k_n(V_{GS} - V_t)^2 = \frac{1}{2}k_n V_{OV}^2 \tag{7.25}$$

where we have neglected channel-length modulation (i.e., we have assumed $\lambda = 0$). Here $V_{OV} = V_{GS} - V_t$ is the overdrive voltage at which the MOSFET is biased to operate. The dc voltage at the drain, V_{DS}, will be

$$V_{DS} = V_{DD} - R_D I_D \tag{7.26}$$

To ensure saturation-region operation, we must have

$$V_{DS} > V_{OV}$$

Furthermore, since the total voltage at the drain will have a signal component superimposed on V_{DS}, V_{DS} has to be sufficiently greater than V_{OV} to allow for the required negative signal swing.

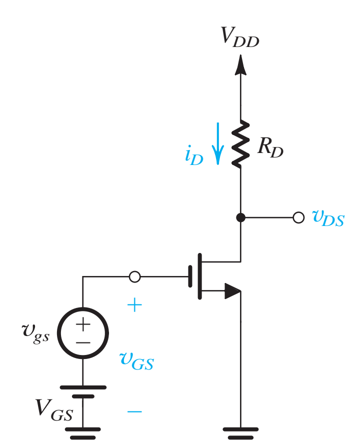


Figure 7.10 Conceptual circuit utilized to study the operation of the MOSFET as a small-signal amplifier.

[3]Practical biasing arrangements will be studied in Section 7.4.

The Signal Current in the Drain Terminal Next, consider the situation with the input signal v_{gs} applied. The total instantaneous gate-to-source voltage will be

$$v_{GS} = V_{GS} + v_{gs} \tag{7.27}$$

resulting in a total instantaneous drain current i_D,

$$\begin{aligned} i_D &= \frac{1}{2}k_n\left(V_{GS} + v_{gs} - V_t\right)^2 \\ &= \frac{1}{2}k_n(V_{GS} - V_t)^2 + k_n(V_{GS} - V_t)v_{gs} + \frac{1}{2}k_n v_{gs}^2 \end{aligned} \tag{7.28}$$

The first term on the right-hand side of Eq. (7.28) can be recognized as the dc bias current I_D (Eq. 7.25). The second term represents a current component that is directly proportional to the input signal v_{gs}. The third term is a current component that is proportional to the square of the input signal. This last component is undesirable because it represents *nonlinear distortion*. To reduce the nonlinear distortion introduced by the MOSFET, the input signal should be kept small so that

$$\frac{1}{2}k_n v_{gs}^2 \ll k_n(V_{GS} - V_t)v_{gs}$$

resulting in

$$v_{gs} \ll 2(V_{GS} - V_t) \tag{7.29}$$

or, equivalently,

$$v_{gs} \ll 2V_{OV} \tag{7.30}$$

If this **small-signal condition** is satisfied, we may neglect the last term in Eq. (7.28) and express i_D as

$$i_D \simeq I_D + i_d \tag{7.31}$$

where

$$i_d = k_n(V_{GS} - V_t)v_{gs}$$

The parameter that relates i_d and v_{gs} is the MOSFET **transconductance** g_m,

$$g_m \equiv \frac{i_d}{v_{gs}} = k_n(V_{GS} - V_t) \tag{7.32}$$

or in terms of the overdrive voltage V_{OV},

$$g_m = k_n V_{OV} \tag{7.33}$$

Figure 7.11 presents a graphical interpretation of the small-signal operation of the MOSFET amplifier. Note that g_m is equal to the slope of the i_D–v_{GS} characteristic at the bias point,

$$g_m \equiv \left.\frac{\partial i_D}{\partial v_{GS}}\right|_{v_{GS}=V_{GS}} \tag{7.34}$$

This is the formal definition of g_m, which can be shown to yield the expressions given in Eqs. (7.32) and (7.33).

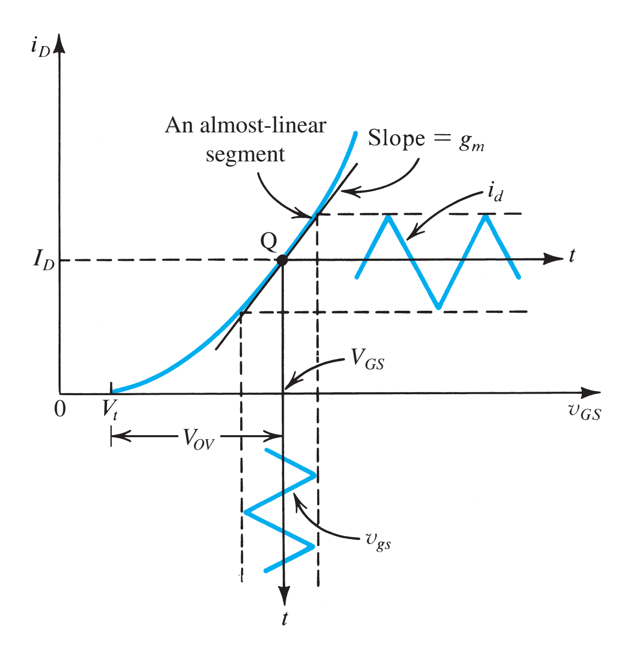


Figure 7.11 Small-signal operation of the MOSFET amplifier.

The Voltage Gain Returning to the circuit of Fig. 7.10, we can express the total instantaneous drain voltage v_{DS} as follows:

$$v_{DS} = V_{DD} - R_D i_D$$

Under the small-signal condition, we have

$$v_{DS} = V_{DD} - R_D(I_D + i_d)$$

which can be rewritten as

$$v_{DS} = V_{DS} - R_D i_d$$

Thus the signal component of the drain voltage is

$$v_{ds} = -i_d R_D = -g_m v_{gs} R_D \tag{7.35}$$

which indicates that the voltage gain is given by

$$A_v \equiv \frac{v_{ds}}{v_{gs}} = -g_m R_D \tag{7.36}$$

The minus sign in Eq. (7.36) indicates that the output signal v_{ds} is 180° out of phase with respect to the input signal v_{gs}. This is illustrated in Fig. 7.12, which shows v_{GS} and v_{DS}. The input signal is assumed to have a triangular waveform with an amplitude much smaller than $2(V_{GS} - V_t)$, the small-signal condition in Eq. (7.29), to ensure linear operation. For operation in the saturation (active) region at all times, the minimum value of v_{DS} should not fall below the corresponding value of v_{GS} by more than V_t. Also, the maximum value of v_{DS} should be

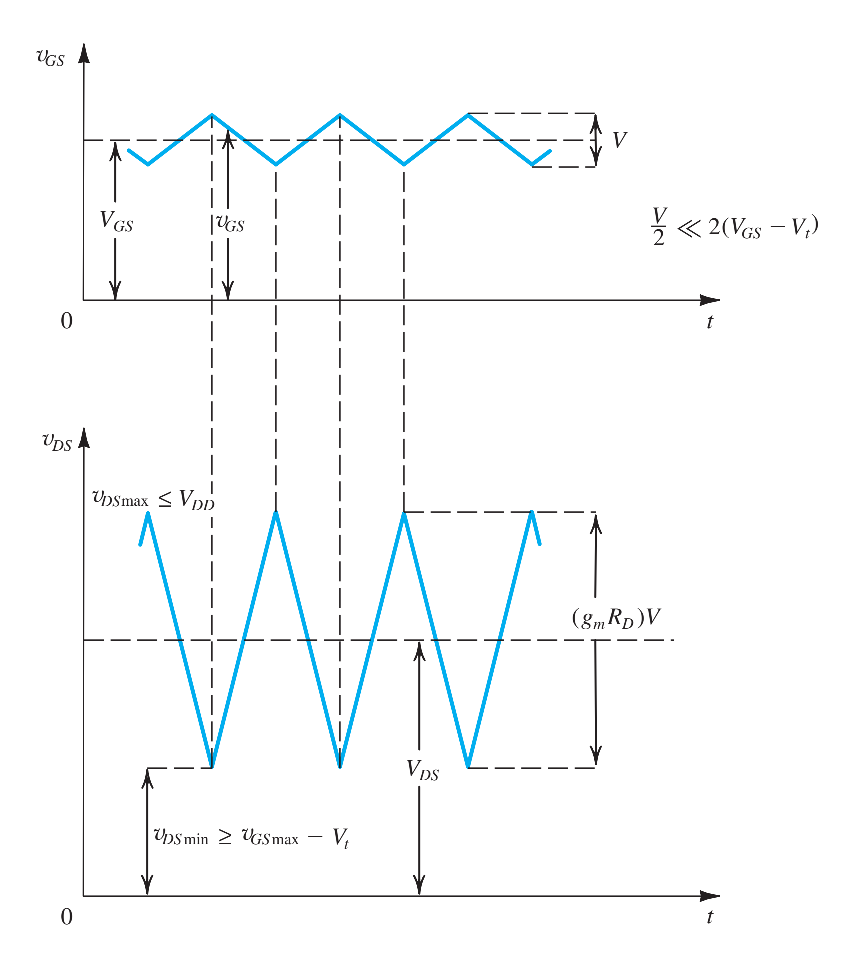


Figure 7.12 Total instantaneous voltages v_{GS} and v_{DS} for the circuit in Fig. 7.10.

smaller than V_{DD}; otherwise the FET will enter the cutoff region and the peaks of the output signal waveform will be clipped off.

Finally, we note that by substituting for g_m from Eq. (7.33) the voltage-gain expression in Eq. (7.36) becomes identical to that derived in Section 7.1—namely, Eq. (7.15).

Separating the DC Analysis and the Signal Analysis From the preceding analysis, we see that under the small-signal approximation, signal quantities are superimposed on dc quantities. For instance, the total drain current i_D equals the dc current I_D plus the signal current i_d, the total drain voltage $v_{DS} = V_{DS} + v_{ds}$, and so on. It follows that the analysis and design can be greatly simplified by separating dc or bias calculations from small-signal calculations. That is, once a stable dc operating point has been established and all dc quantities calculated, we may then perform signal analysis ignoring dc quantities.

Small-Signal Equivalent-Circuit Models From a signal point of view, the FET behaves as a voltage-controlled current source. It accepts a signal v_{gs} between gate and source and provides a current $g_m v_{gs}$ at the drain terminal. The input resistance of this controlled source is very high—ideally, infinite. The output resistance—that is, the resistance looking into the

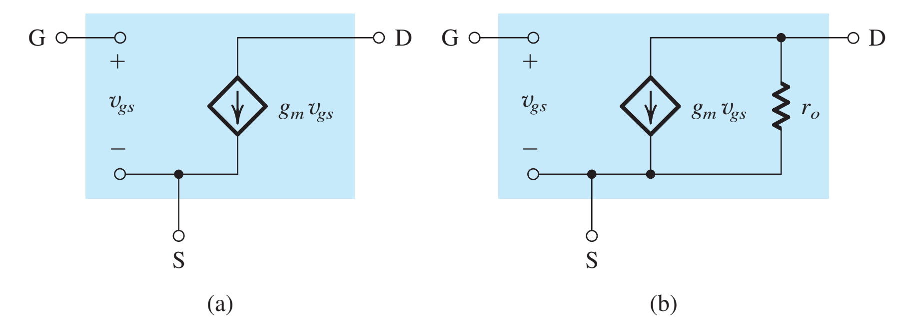


Figure 7.13 Small-signal models for the MOSFET: **(a)** neglecting the dependence of i_D on v_{DS} in the active region (the channel-length modulation effect) and **(b)** including the effect of channel-length modulation, modeled by output resistance $r_o = |V_A|/I_D$. These models apply equally well for both NMOS and PMOS transistors.

drain—also is high, and we have assumed it to be infinite thus far. Putting all of this together, we arrive at the circuit in Fig. 7.13(a), which represents the small-signal operation of the MOSFET and is thus a **small-signal model** or a **small-signal equivalent circuit**.

In the analysis of a MOSFET amplifier circuit, the transistor can be replaced by the equivalent-circuit model shown in Fig. 7.13(a). The rest of the circuit remains unchanged except that *ideal constant dc voltage sources are replaced by short circuits*. This is a result of the fact that the voltage across an ideal constant dc voltage source does not change, and thus there will always be a zero voltage signal across a constant dc voltage source. A dual statement applies for constant dc current sources; namely, the signal current of an ideal constant dc current source will always be zero, and thus *an ideal constant dc current source can be replaced by an open circuit* in the small-signal equivalent circuit of the amplifier. The circuit resulting can then be used to perform any required signal analysis, such as calculating voltage gain.

The most serious shortcoming of the small-signal model of Fig. 7.13(a) is that it assumes the drain current in saturation to be independent of the drain voltage. From our study of the MOSFET characteristics in saturation, we know that the drain current does in fact depend on v_{DS} in a linear manner. Such dependence was modeled by a finite resistance r_o between drain and source, whose value was given by Eq. (5.27) in Section 5.2.4, which we repeat here (with the prime on I_D dropped) as

$$r_o = \frac{|V_A|}{I_D} \tag{7.37}$$

where $V_A = 1/\lambda$ is a MOSFET parameter that either is specified or can be measured. It should be recalled that for a given process technology, V_A is proportional to the MOSFET channel length. The current I_D is the value of the dc drain current without the channel-length modulation taken into account; that is,

$$I_D = \frac{1}{2} k_n V_{OV}^2 \tag{7.38}$$

Typically, r_o is in the range of 10 kΩ to 1000 kΩ. It follows that the accuracy of the small-signal model can be improved by including r_o in parallel with the controlled source, as shown in Fig. 7.13(b).

It is important to note that the small-signal model parameters g_m and r_o depend on the dc bias point of the MOSFET.

Returning to the amplifier of Fig. 7.10, we find that replacing the MOSFET with the small-signal model of Fig. 7.13(b) results in the voltage-gain expression

$$A_v = \frac{v_{ds}}{v_{gs}} = -g_m(R_D \parallel r_o) \tag{7.39}$$

Thus, the finite output resistance r_o results in a reduction in the magnitude of the voltage gain.

Although the analysis above is performed on an NMOS transistor, the results, and the equivalent-circuit models of Fig. 7.13, apply equally well to PMOS devices, except for using $|V_{GS}|, |V_t|, |V_{OV}|$, and $|V_A|$ and replacing k_n with k_p.

The Transconductance g_m We shall now take a closer look at the MOSFET transconductance given by Eq. (7.32), which we rewrite with $k_n = k_n'(W/L)$ as follows:

$$g_m = k_n'(W/L)(V_{GS} - V_t) = k_n'(W/L)V_{OV} \tag{7.40}$$

This relationship indicates that g_m is proportional to the process transconductance parameter $k_n' = \mu_n C_{ox}$ and to the W/L ratio of the MOS transistor; hence to obtain relatively large transconductance the device must be short and wide. We also observe that for a given device the transconductance is proportional to the overdrive voltage, $V_{OV} = V_{GS} - V_t$, the amount by which the bias voltage V_{GS} exceeds the threshold voltage V_t. Note, however, that increasing g_m by biasing the device at a larger V_{GS} has the disadvantage of reducing the allowable voltage signal swing at the drain.

Another useful expression for g_m can be obtained by substituting for V_{OV} in Eq. (7.40) by $\sqrt{2I_D/(k_n'(W/L))}$ [from Eq. (7.25)]:

$$g_m = \sqrt{2k_n'}\sqrt{W/L}\sqrt{I_D} \tag{7.41}$$

This expression shows two things:

1. For a given MOSFET, g_m is proportional to the square root of the dc bias current.
2. At a given bias current, g_m is proportional to $\sqrt{W/L}$.

In contrast, as we shall see shortly, the transconductance of the bipolar junction transistor (BJT) is proportional to the bias current and is independent of the physical size and geometry of the device.

To gain some insight into the values of g_m obtained in MOSFETs consider an integrated-circuit device operating at $I_D = 0.5$ mA and having $k_n' = 120\ \mu\text{A/V}^2$. Equation (7.41) shows that for $W/L = 1, g_m = 0.35$ mA/V, whereas a device for which $W/L = 100$ has $g_m = 3.5$ mA/V. In contrast, a BJT operating at a collector current of 0.5 mA has $g_m = 20$ mA/V.

Yet another useful expression for g_m of the MOSFET can be obtained by substituting for $k_n'(W/L)$ in Eq. (7.40) by $2I_D/(V_{GS} - V_t)^2$:

$$g_m = \frac{2I_D}{V_{GS} - V_t} = \frac{2I_D}{V_{OV}} \tag{7.42}$$

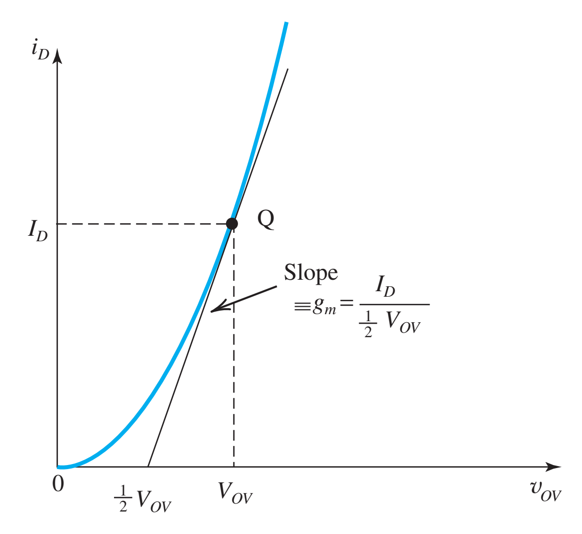


Figure 7.14 The slope of the tangent at the bias point Q intersects the v_{OV} axis at $\frac{1}{2}V_{OV}$. Thus, $g_m = I_D/(\frac{1}{2}V_{OV})$.

A convenient graphical construction that clearly illustrates this relationship is shown in Fig. 7.14.

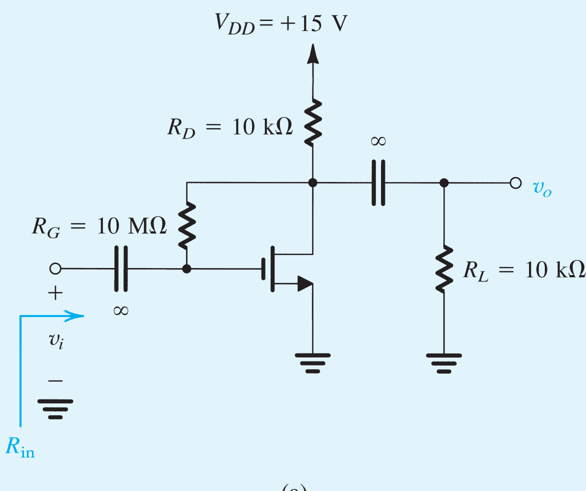
In summary, there are three different relationships for determining g_m—Eqs. (7.40), (7.41), and (7.42)—and there are three design parameters—(W/L), V_{OV}, and I_D, any two of which can be chosen independently. That is, the designer may choose to operate the MOSFET with a certain overdrive voltage V_{OV} and at a particular current I_D; the required W/L ratio can then be found and the resulting g_m determined.[4]

Example 7.3

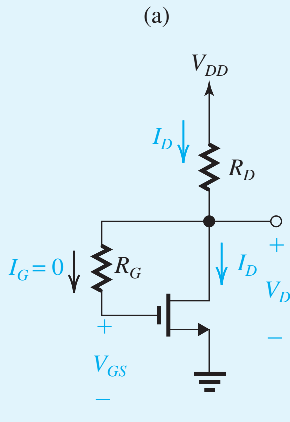
Figure 7.15(a) shows a discrete MOSFET amplifier utilizing a drain-to-gate resistance R_G for biasing purposes. Such a biasing arrangement will be studied in Section 7.4. The input signal v_i is coupled to the gate via a large capacitor, and the output signal at the drain is coupled to the load resistance R_L via another large capacitor. We wish to analyze this amplifier circuit to determine its small-signal voltage gain, its input resistance, and the largest allowable input signal. The transistor has $V_t = 1.5$ V, $k_n'(W/L) = 0.25$ mA/V^2, and $V_A = 50$ V. Assume the coupling capacitors to be sufficiently large so as to act as short circuits at the signal frequencies of interest.

[4]This assumes that the circuit designer is also designing the device, as is typically the case in IC design. On the other hand, a circuit designer working with a discrete-circuit MOSFET obviously does not have the freedom to change its W/L ratio. Thus, in this case there are only two design parameters—V_{OV} and I_D, and only one can be specified by the designer.

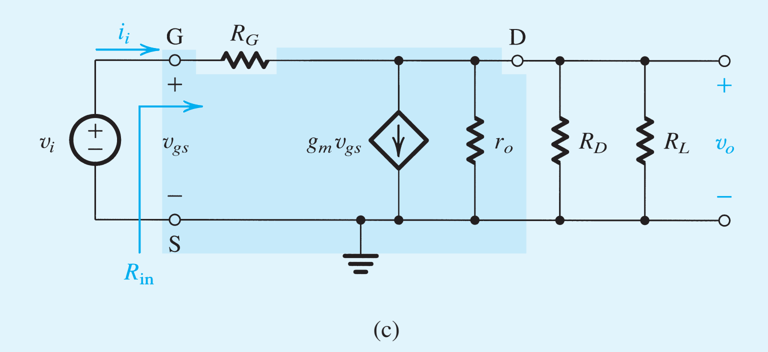
Example 7.3 *continued*



(a)

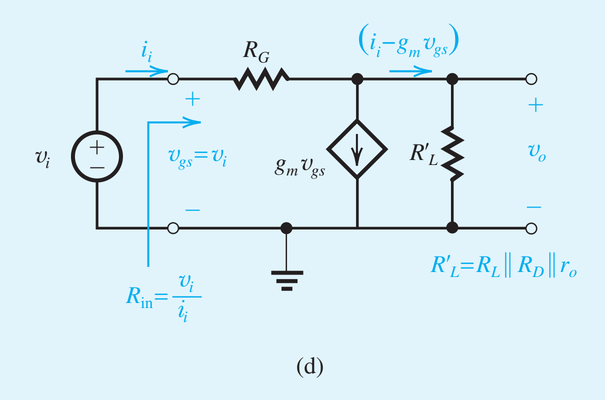


(b)



(c)

Figure 7.15 Example 7.3: **(a)** amplifier circuit; **(b)** circuit for determining the dc operating point; **(c)** the amplifier small-signal equivalent circuit; **(d)** a simplified version of the circuit in **(c)**.



(d)

Figure 7.15 *continued*

Solution

We first determine the dc operating point. For this purpose, we eliminate the input signal v_i, and open-circuit the two coupling capacitors (since they block dc currents). The result is the circuit shown in Fig. 7.14(b). We note that since $I_G = 0$, the dc voltage drop across R_G will be zero, and

$$V_{GS} = V_{DS} = V_{DD} - R_D I_D \tag{7.43}$$

With $V_{DS} = V_{GS}$, the NMOS transistor will be operating in saturation. Thus,

$$I_D = \frac{1}{2} k_n (V_{GS} - V_t)^2 \tag{7.44}$$

where, for simplicity, we have neglected the effect of channel-length modulation on the dc operating point. Substituting $V_{DD} = 15$ V, $R_D = 10$ kΩ, $k_n = 0.25$ mA/V^2, and $V_t = 1.5$ V in Eqs. (7.43) and (7.44), and substituting for V_{GS} from Eq. (7.43) into Eq. (7.44) results in a quadratic equation in I_D. Solving the latter and discarding the root that is not physically meaningful yields the solution

$$I_D = 1.06 \text{ mA}$$

which corresponds to

$$V_{GS} = V_{DS} = 4.4 \text{ V}$$

and

$$V_{OV} = 4.4 - 1.5 = 2.9 \text{ V}$$

Next we proceed with the small-signal analysis of the amplifier. Toward that end we replace the MOSFET with its small-signal model to obtain the small-signal equivalent circuit of the amplifier, shown in Fig. 7.15(c). Observe that we have replaced the coupling capacitors with short circuits. The dc voltage supply V_{DD} has also been replaced with a short circuit to ground.

Example 7.3 *continued*

The values of the transistor small-signal parameters g_m and r_o can be determined by using the dc bias quantities found above, as follows:

$$g_m = k_n V_{OV}$$
$$= 0.25 \times 2.9 = 0.725 \text{ mA/V}$$
$$r_o = \frac{V_A}{I_D} = \frac{50}{1.06} = 47 \text{ k}\Omega$$

Next we use the equivalent circuit of Fig. 7.15(c) to determine the input resistance $R_{\text{in}} \equiv v_i/i_i$ and the voltage gain $A_v = v_o/v_i$. Toward that end we simplify the circuit by combining the three parallel resistances r_o, R_D, and R_L in a single resistance R'_L,

$$R'_L = R_L || R_D || r_o$$
$$= 10||10||47 = 4.52 \text{ k}\Omega$$

as shown in Fig. 7.15(d). For the latter circuit we can write the two equations

$$v_o = (i_i - g_m v_{gs})R'_L \tag{7.45}$$

and

$$i_i = \frac{v_{gs} - v_o}{R_G} \tag{7.46}$$

Substituting for i_i from Eq. (7.46) into Eq. (7.45) results in the following expression for the voltage gain $A_v \equiv v_o/v_i = v_o/v_{gs}$:

$$A_v = -g_m R'_L \frac{1 - (1/g_m R_G)}{1 + (R'_L/R_G)}$$

Since R_G is very large, $g_m R_G \gg 1$ and $R'_L/R_G \ll 1$ (the reader can easily verify this), and the gain expression can be approximated as

$$A_v \simeq -g_m R'_L \tag{7.47}$$

Substituting $g_m = 0.725$ mA/V and $R'_L = 4.52$ kΩ yields

$$A_v = -3.3 \text{ V/V}$$

To obtain the input resistance, we substitute in Eq. (7.46) for $v_o = A_v v_{gs} = -g_m R'_L v_{gs}$, then use $R_{\text{in}} \equiv v_i/i_i = v_{gs}/i_i$ to obtain

$$R_{\text{in}} = \frac{R_G}{1 + g_m R'_L} \tag{7.48}$$

This is an interesting relationship: The input resistance decreases as the gain $(g_m R'_L)$ is increased. The value of R_{in} can now be determined; it is

$$R_{in} = \frac{10\text{ M}\Omega}{1+3.3} = 2.33\text{ M}\Omega$$

which is still very large.

The largest allowable input signal $\hat{v}_i$ is constrained by the need to keep the transistor in saturation at all times; that is,

$$v_{DS} \geq v_{GS} - V_t$$

Enforcing this condition with equality at the point v_{GS} is maximum and v_{DS} is minimum, we write

$$v_{DS\min} = v_{GS\max} - V_t$$

$$V_{DS} - |A_v|\hat{v}_i = V_{GS} + \hat{v}_i - V_t$$

Since $V_{DS} = V_{GS}$, we obtain

$$\hat{v}_i = \frac{V_t}{|A_v|+1}$$

This is a general relationship that applies to this circuit irrespective of the component values. Observe that it simply states that the maximum signal swing is determined by the fact that the bias arrangement makes $V_D = V_G$ and thus, to keep the MOSFET out of the triode region, the signal between D and G is constrained to be equal to V_t. For our particular design,

$$\hat{v}_i = \frac{1.5}{3.3+1} = 0.35\text{ V}$$

Since $V_{OV} = 2.9$ V, a v_i of 0.35 is indeed much smaller than $2V_{OV} = 5.8$ V; thus the assumption of linear operation is justified.

A modification of this circuit that increases the allowable signal swing is investigated in Problem 7.103.

EXERCISE

D7.4 Consider the amplifier circuit of Fig. 7.15(a) without the load resistance R_L and with channel-length modulation neglected. Let $V_{DD} = 5$ V, $V_t = 0.7$ V, and $k_n = 1\text{ mA/V}^2$. Find V_{OV}, I_D, R_D, and R_G to obtain a voltage gain of −25 V/V and an input resistance of 0.5 M Ω. What is the maximum allowable input signal, $\hat{v}_i$?

Ans. 0.319 V; 50.9 μA; 78.5 kΩ; 13 MΩ; 27 mV

The T Equivalent-Circuit Model Through a simple circuit transformation it is possible to develop an alternative equivalent-circuit model for the MOSFET. The development of such a model, known as the T model, is illustrated in Fig. 7.16. Figure 7.16(a) shows the equivalent circuit studied above without r_o. In Fig. 7.16(b) we have added a second $g_m v_{gs}$ current source in series with the original controlled source. This addition obviously does not change the terminal currents and is thus allowed. The newly created circuit node, labeled X, is joined to the gate terminal G in Fig. 7.16(c). Observe that the gate current does not change—that is, it remains equal to zero—and thus this connection does not alter the terminal characteristics. We now note that we have a controlled current source $g_m v_{gs}$ connected across its control voltage v_{gs}. We can replace this controlled source by a resistance as long as this resistance draws an equal current as the source. (See the source-absorption theorem in Appendix D.) Thus the value of the resistance is $v_{gs}/g_m v_{gs} = 1/g_m$. This replacement is shown in Fig. 7.16(d), which depicts

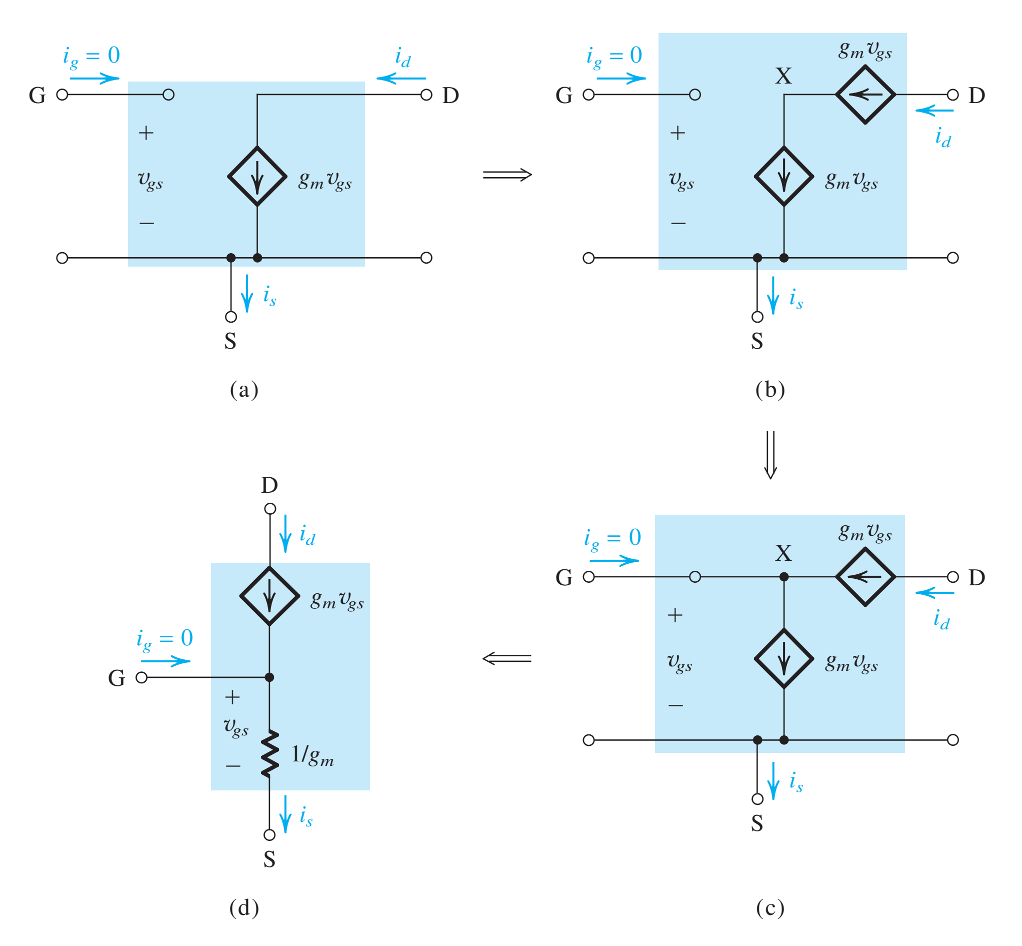


Figure 7.16 Development of the T equivalent-circuit model for the MOSFET. For simplicity, r_o has been omitted; however, it may be added between D and S in the T model of **(d)**.

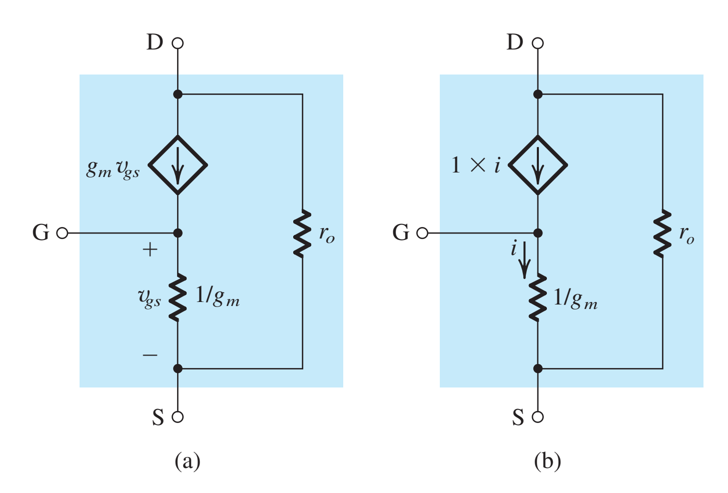


Figure 7.17 **(a)** The T model of the MOSFET augmented with the drain-to-source resistance r_o. **(b)** An alternative representation of the T model.

the alternative model. Observe that i_g is still zero, $i_d = g_m v_{gs}$, and $i_s = v_{gs}/(1/g_m) = g_m v_{gs}$, all the same as in the original model in Fig. 7.16(a).

The model of Fig. 7.16(d) shows that the resistance between gate and source looking into the source is $1/g_m$. This observation and the T model prove useful in many applications. Note that the resistance between gate and source, looking into the gate, is infinite.

In developing the T model we did not include r_o. If desired, this can be done by incorporating in the circuit of Fig. 7.16(d) a resistance r_o between drain and source, as shown in Fig. 7.17(a). An alternative representation of the T model, in which the voltage-controlled current source is replaced with a current-controlled current source, is shown in Fig. 7.17(b).

Finally, we should note that in order to distinguish the model of Fig. 7.13(b) from the equivalent T model, the former is sometimes referred to as the **hybrid-π model**, a carryover from the bipolar transistor literature. The origin of this name will be explained shortly.

Example 7.4

Figure 7.18(a) shows a MOSFET amplifier biased by a constant-current source I. Assume that the values of I and R_D are such that the MOSFET operates in the saturation region. The input signal v_i is coupled to the source terminal by utilizing a large capacitor C_{C1}. Similarly, the output signal at the drain is taken through a large coupling capacitor C_{C2}. Find the input resistance R_{in} and the voltage gain v_o/v_i. Neglect channel-length modulation.

Example 7.4 *continued*

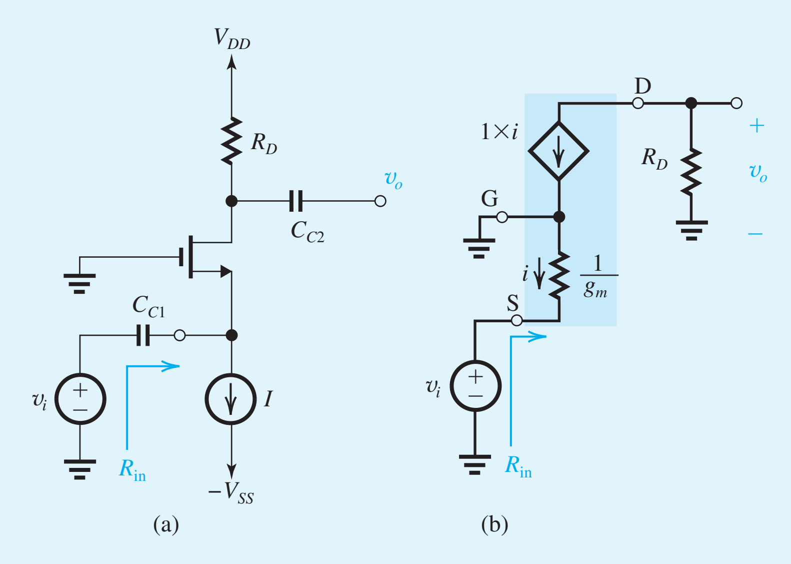


Figure 7.18 **(a)** Amplifier circuit for Example 7.4. **(b)** Small-signal equivalent circuit of the amplifier in **(a)**.

Solution

Replacing the MOSFET with its T equivalent-circuit model results in the amplifier equivalent circuit shown in Fig. 7.18(b). Observe that the dc current source I is replaced with an open circuit and the dc voltage source V_{DD} is replaced by a short circuit. The large coupling capacitors have been replaced by short circuits. From the equivalent-circuit model we determine

$$R_{in} = \frac{v_i}{-i} = 1/g_m$$

and

$$v_o = -iR_D = \left(\frac{v_i}{1/g_m}\right)R_D = g_m R_D v_i$$

Thus,

$$A_v \equiv \frac{v_o}{v_i} = g_m R_D$$

We note that this amplifier, known as the common-gate amplifier because the gate at ground potential is common to both the input and output ports, has a low input resistance $(1/g_m)$ and a noninverting gain. We shall study this amplifier type in Section 7.3.5.

EXERCISE

7.5 Use the T model of Fig. 7.17(b) to show that a MOSFET whose drain is connected to its gate exhibits an incremental resistance equal to $[(1/g_m) \| r_o]$.
Ans. See Fig. E7.5.

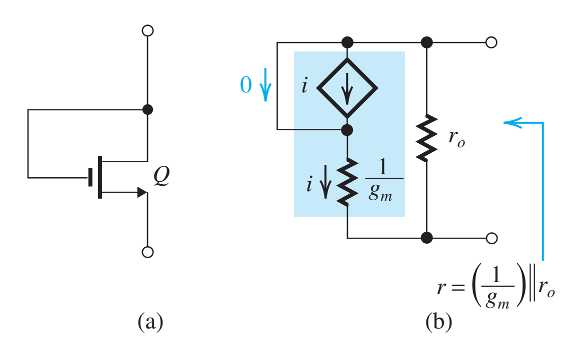


Figure E7.5 Circuits for Exercise 7.5. Note that the bias arrangement of Q is not shown.

Modeling the Body Effect As mentioned earlier (see Section 5.4), the body effect occurs in a MOSFET when the source is not tied to the substrate (which is always connected to the most negative power supply in the integrated circuit for n-channel devices and to the most positive for p-channel devices). Thus the substrate (body) will be at signal ground, but since the source is not, a signal voltage v_{bs} develops between the body (B) and the source (S). The substrate then acts as a "second gate" or a **backgate** for the MOSFET. Thus the signal v_{bs} gives rise to a drain-current component, which we shall write as $g_{mb}v_{bs}$, where g_{mb} is the **body transconductance**, defined as

$$g_{mb} \equiv \left.\frac{\partial i_D}{\partial v_{BS}}\right|_{\substack{v_{GS}=\text{constant}\\ v_{DS}=\text{constant}}} \tag{7.49}$$

Recalling that i_D depends on v_{BS} through the dependence of V_t on V_{BS}, we can show that

$$g_{mb} = \chi g_m \tag{7.50}$$

where

$$\chi \equiv \frac{\partial V_t}{\partial V_{SB}} = \frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} \tag{7.51}$$

Typically the value of χ lies in the range 0.1 to 0.3.

Figure 7.19 shows the MOSFET model augmented to include the controlled source $g_{mb}v_{bs}$ that models the body effect. Ideally, this is the model to be used whenever the source is not connected to the substrate. It has been found, however, that except in some very particular

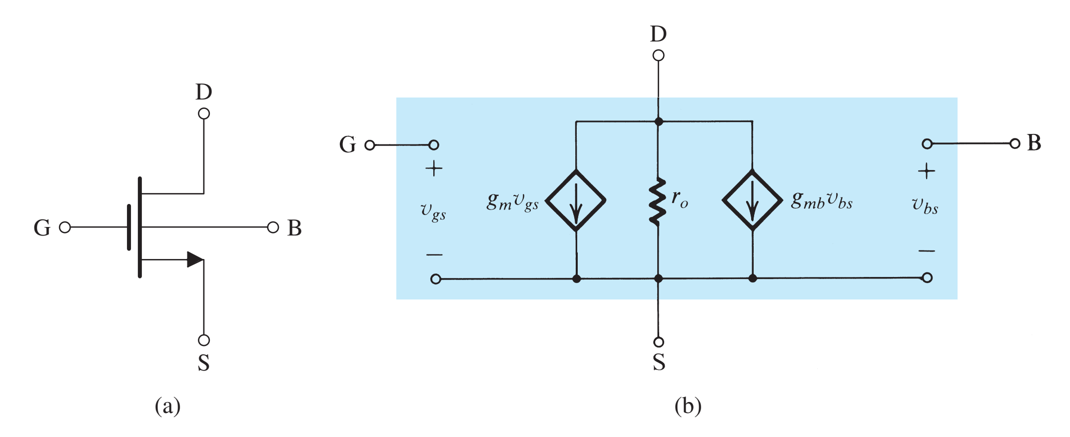


Figure 7.19 Small-signal, equivalent-circuit model of a MOSFET in which the source is not connected to the body.

situations, the body effect can generally be ignored in the initial, pencil-and-paper design of MOSFET amplifiers.

Finally, although the analysis above was performed on an NMOS transistor, the results and the equivalent circuit of Fig. 7.19 apply equally well to PMOS transistors, except for using $|V_{GS}|$, $|V_t|$, $|V_{OV}|$, $|V_A|$, $|V_{SB}|$, $|\gamma|$, and $|\lambda|$ and replacing k_n' with k_p' in the appropriate formula.

EXERCISES

7.6 For the amplifier in Fig. 7.4, let $V_{DD} = 5$ V, $R_D = 10$ kΩ, $V_t = 1$ V, $k_n' = 20\ \mu\text{A/V}^2$, $W/L = 20$, $V_{GS} = 2$ V, and $\lambda = 0$.
(a) Find the dc current I_D and the dc voltage V_{DS}.
(b) Find g_m.
(c) Find the voltage gain.
(d) If $v_{gs} = 0.2 \sin \omega t$ volts, find v_{ds} assuming that the small-signal approximation holds. What are the minimum and maximum values of v_{DS}?
(e) Use Eq. (7.28) to determine the various components of i_D. Using the identity ($\sin^2 \omega t = \frac{1}{2} - \frac{1}{2}\cos 2\,\omega t$), show that there is a slight shift in I_D (by how much?) and that there is a second-harmonic component (i.e., a component with frequency $2\,\omega$). Express the amplitude of the second-harmonic component as a percentage of the amplitude of the fundamental. (This value is known as the second-harmonic distortion.)
Ans. (a) 200 μA, 3 V; (b) 0.4 mA/V; (c) –4 V/V; (d) $v_{ds} = -0.8 \sin \omega t$ volts, 2.2 V, 3.8 V; (e) $i_D = (204 + 80\ \sin \omega t - 4 \cos 2\,\omega t)$ μA, 5%

7.7 An NMOS transistor has $\mu_n C_{ox} = 60\ \mu\text{A/V}^2$, $W/L = 40$, $V_t = 1$ V, and $V_A = 15$ V. Find g_m and r_o when (a) the bias voltage $V_{GS} = 1.5$ V, (b) the bias current $I_D = 0.5$ mA.
Ans. (a) 1.2 mA/V, 50 kΩ; (b) 1.55 mA/V, 30 kΩ

7.8 A MOSFET is to operate at $I_D = 0.1$ mA and is to have $g_m = 1$ mA/V. If $k_n' = 50\ \mu A/V^2$, find the required W/L ratio and the overdrive voltage.
Ans. 100; 0.2 V

7.9 For a fabrication process for which $\mu_p \simeq 0.4\mu_n$, find the ratio of the width of a PMOS transistor to the width of an NMOS transistor so that the two devices have equal g_m for the same bias conditions. The two devices have equal channel lengths.
Ans. 2.5

7.10 A PMOS transistor has $V_t = -1$ V, $k_p' = 60\ \mu A/V^2$, and $W/L = 16\ \mu m/0.8\ \mu m$. Find I_D and g_m when the device is biased at $V_{GS} = -1.6$ V. Also, find the value of r_o if λ (at $L = 1\ \mu m$) = –0.04 V^{-1}.
Ans. 216 μA; 0.72 mA/V; 92.6 kΩ

7.11 Derive an expression for $(g_m r_o)$ in terms of V_A and V_{OV}. As we shall see in Chapter 8, this is an important transistor parameter and is known as the intrinsic gain. Evaluate the value of $g_m r_o$ for an NMOS transistor fabricated in a 0.8-μm CMOS process for which $V_A' = 12.5$ V/μm of channel length. Let the device have minimum channel length and be operated at an overdrive voltage of 0.2 V.
Ans. $g_m r_o = 2V_A/V_{OV}$; 100 V/V

7.2.2 The BJT Case

We next consider the small-signal operation of the BJT and develop small-signal equivalent-circuit models that represent its operation at a given bias point. The following development parallels what we used for the MOSFET except that here we have an added complication: The BJT draws a finite base current. As will be seen shortly, this phenomenon (finite β) manifests itself as a finite input resistance looking into the base of the BJT (as compared to the infinite input resistance looking into the gate of the MOSFET).

Consider the *conceptual* amplifier circuit shown in Fig. 7.20(a). Here the base–emitter junction is forward biased by a dc voltage V_{BE}. The reverse bias of the collector–base junction is established by connecting the collector to another power supply of voltage V_{CC} through a resistor R_C. The input signal to be amplified is represented by the voltage source v_{be} that is superimposed on V_{BE}.

The DC Bias Point We consider first the dc bias conditions by setting the signal v_{be} to zero. The circuit reduces to that in Fig. 7.20(b), and we can write the following relationships for the dc currents and voltages:

$$I_C = I_S e^{V_{BE}/V_T} \tag{7.52}$$

$$I_E = I_C/\alpha \tag{7.53}$$

$$I_B = I_C/\beta \tag{7.54}$$

$$V_{CE} = V_{CC} - I_C R_C \tag{7.55}$$

For active-mode operation, V_{CE} should be greater than $(V_{BE} - 0.4)$ by an amount that allows for the required negative signal swing at the collector.

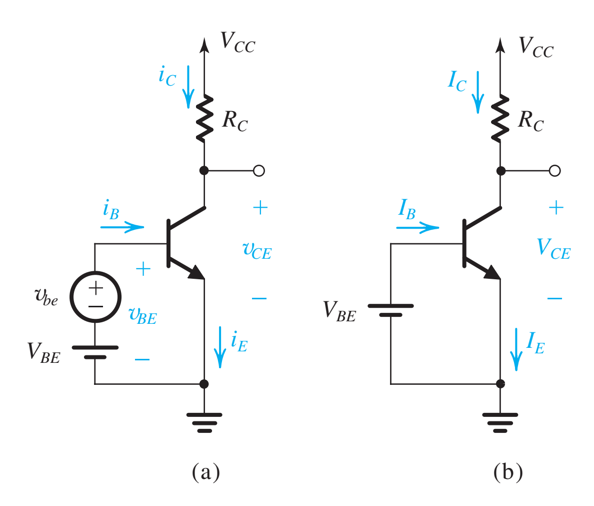


Figure 7.20 (**a**) Conceptual circuit to illustrate the operation of the transistor as an amplifier. (**b**) The circuit of (**a**) with the signal source v_{be} eliminated for dc (bias) analysis.

The Collector Current and the Transconductance If a signal v_{be} is applied as shown in Fig. 7.20(a), the total instantaneous base–emitter voltage v_{BE} becomes

$$v_{BE} = V_{BE} + v_{be}$$

Correspondingly, the collector current becomes

$$\begin{aligned} i_C &= I_S e^{v_{BE}/V_T} = I_S e^{(V_{BE}+v_{be})/V_T} \\ &= I_S e^{V_{BE}/V_T} e^{v_{be}/V_T} \end{aligned}$$

Use of Eq. (7.52) yields

$$i_C = I_C e^{v_{be}/V_T} \tag{7.56}$$

Now, if $v_{be} \ll V_T$, we may approximate Eq. (7.56) as

$$i_C \simeq I_C \left(1 + \frac{v_{be}}{V_T}\right) \tag{7.57}$$

Here we have expanded the exponential in Eq. (7.56) in a series and retained only the first two terms. That is, we have assumed that

$$v_{be} \ll V_T \tag{7.58}$$

so that we can neglect the higher-order terms in the exponential series expansion. The condition in Eq. (7.58) is the **small-signal approximation** for the BJT and corresponds to that in Eq. (7.29) for the MOSFET case. The small-signal approximation for the BJT is valid only for v_{be} less than 5 mV or 10 mV, at most. Under this approximation, the total collector current is given by Eq. (7.57) and can be rewritten

$$i_C = I_C + \frac{I_C}{V_T} v_{be} \tag{7.59}$$

Thus the collector current is composed of the dc bias value I_C and a signal component i_c,

$$i_c = \frac{I_C}{V_T} v_{be} \tag{7.60}$$

This equation relates the signal current in the collector to the corresponding base–emitter signal voltage. It can be rewritten as

$$i_c = g_m v_{be} \tag{7.61}$$

where g_m is the **transconductance**, and from Eq. (7.60), it is given by

$$g_m = \frac{I_C}{V_T} \tag{7.62}$$

We observe that the transconductance of the BJT is directly proportional to the collector bias current I_C. Thus to obtain a constant predictable value for g_m, we need a constant predictable I_C. Also, we note that BJTs have relatively high transconductance in comparison to MOSFETs: for instance, at $I_C = 1$ mA, $g_m \simeq 40$ mA/V. Finally, unlike the MOSFET, whose g_m depends on the device dimensions (W and L), g_m of a BJT depends only on the dc collector current at which it is biased to operate.

A graphical interpretation for g_m is given in Fig. 7.21, where it is shown that g_m is equal to the slope of the tangent to the i_C–v_{BE} characteristic curve at $i_C = I_C$ (i.e., at the bias point Q). Thus,

$$g_m = \left.\frac{\partial i_C}{\partial v_{BE}}\right|_{i_C = I_C} \tag{7.63}$$

The small-signal approximation implies keeping the signal amplitude sufficiently small that *operation is restricted to an almost-linear segment of the* i_C–v_{BE} *exponential curve*. Increasing the signal amplitude will result in the collector current having components nonlinearly related to v_{be}.

EXERCISES

7.12 Use Eq. (7.63) to derive the expression for g_m in Eq. (7.62).

7.13 Calculate the value of g_m for a BJT biased at $I_C = 0.5$ mA.
Ans. 20 mA/V

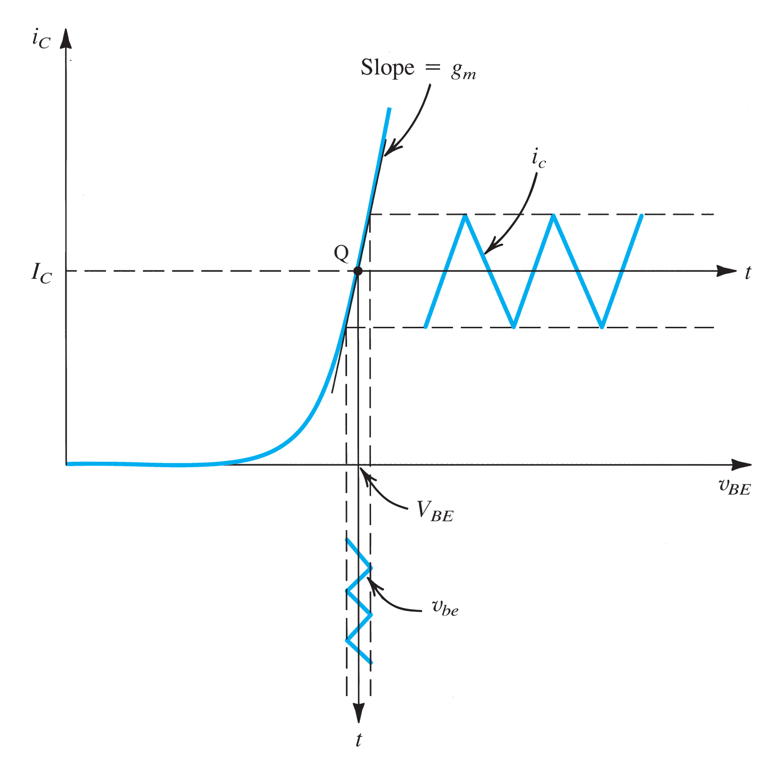


Figure 7.21 Linear operation of the transistor under the small-signal condition: A small-signal v_{be} with a triangular waveform is superimposed on the dc voltage V_{BE}. It gives rise to a collector-signal current i_c, also of triangular waveform, superimposed on the dc current I_C. Here, $i_c = g_m v_{be}$, where g_m is the slope of the i_C–v_{BE} curve at the bias point Q.

The Base Current and the Input Resistance at the Base To determine the resistance seen by v_{be}, we first evaluate the total base current i_B using Eq. (7.59), as follows:

$$i_B = \frac{i_C}{\beta} = \frac{I_C}{\beta} + \frac{1}{\beta}\frac{I_C}{V_T}v_{be}$$

Thus,

$$i_B = I_B + i_b \tag{7.64}$$

where I_B is equal to I_C/β and the signal component i_b is given by

$$i_b = \frac{1}{\beta}\frac{I_C}{V_T}v_{be} \tag{7.65}$$

Substituting for I_C/V_T by g_m gives

$$i_b = \frac{g_m}{\beta}v_{be} \tag{7.66}$$

The small-signal input resistance between base and emitter, *looking into the base,* is denoted by r_π and is defined as

$$r_\pi \equiv \frac{v_{be}}{i_b} \tag{7.67}$$

Using Eq. (7.66) gives

$$r_\pi = \frac{\beta}{g_m} \tag{7.68}$$

Thus r_π is directly dependent on β and is inversely proportional to the bias current I_C. Substituting for g_m in Eq. (7.68) from Eq. (7.62) and replacing I_C/β by I_B gives an alternative expression for r_π,

$$r_\pi = \frac{V_T}{I_B} \tag{7.69}$$

Here, we recall that because the gate current of the MOSFET is zero (at dc and low frequencies) the input resistance at the gate is infinite; that is, in the MOSFET there is no counterpart to r_π.[5]

EXERCISE

7.14 A BJT amplifier is biased to operate at a constant collector current $I_C = 0.5$ mA irrespective of the value β. If the transistor manufacturer specifies β to range from 50 to 200, give the expected range of g_m, I_B, and r_π.
Ans. g_m is constant at 20 mA/V; $I_B = 10$ μA to 2.5 μA; $r_\pi = 2.5$ kΩ to 10 kΩ

The Emitter Current and the Input Resistance at the Emitter The total emitter current i_E can be determined using Eq. (7.59) as

$$i_E = \frac{i_C}{\alpha} = \frac{I_C}{\alpha} + \frac{i_c}{\alpha}$$

Thus,

$$i_E = I_E + i_e \tag{7.70}$$

where I_E is equal to I_C/α and the signal current i_e is given by

$$i_e = \frac{i_c}{\alpha} = \frac{I_C}{\alpha V_T} v_{be} = \frac{I_E}{V_T} v_{be} \tag{7.71}$$

[5]At high frequencies, the input capacitance at the MOSFET gate makes the input current finite (see Chapter 10).

If we denote the small-signal resistance between base and emitter *looking into the emitter* by r_e, it can be defined as

$$r_e \equiv \frac{v_{be}}{i_e} \tag{7.72}$$

Using Eq. (7.71) we find that r_e, called the **emitter resistance**, is given by

$$r_e = \frac{V_T}{I_E} \tag{7.73}$$

Comparison with Eq. (7.62) reveals that

$$r_e = \frac{\alpha}{g_m} \simeq \frac{1}{g_m} \tag{7.74}$$

The relationship between r_π and r_e can be found by combining their respective definitions in Eqs. (7.67) and (7.72) as

$$v_{be} = i_b r_\pi = i_e r_e$$

Thus,

$$r_\pi = (i_e/i_b) r_e$$

which yields

$$r_\pi = (\beta + 1) r_e \tag{7.75}$$

Figure 7.22 illustrates the definition of r_π and r_e.

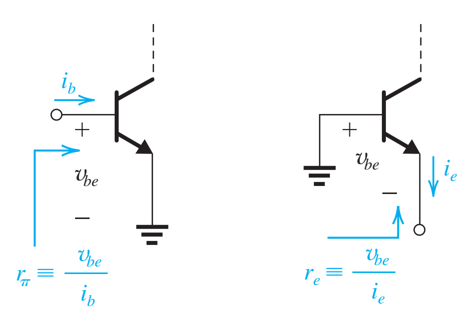


Figure 7.22 Illustrating the definition of r_π and r_e.

Finally, a comparison with the MOSFET would be useful: For the MOSFET, $\alpha = 1$ and the resistance looking into the source is simply $1/g_m$.

SHOCKLEY AND SILICON VALLEY:

In 1956 William Bradford Shockley started a new company, Shockley Semiconductor Laboratory in Mountain View, California (near Stanford, his birthplace). While at Bell Labs, together with John Bardeen and Walter Brattain, he had invented the BJT. At Shockley, the initial concentration was on developing semiconductor devices, particularly a new four-layer diode. But Shockley's scientific genius and ability to select and attract good team members, first demonstrated at Bell Labs, was not accompanied by comparable talent for management. Consequently, in 1957, eight of his team members (the so-called Traitorous Eight, including Gordon Moore and Robert Noyce) left Shockley to create Fairchild Semiconductor. It was a propitious time: The first *Sputnik* was launched a month later, and the ensuing space race accelerated demand for solid-state circuits. Decades passed, and in 2002 a group of some 30 individuals who had been associated with Silicon Valley since 1956 met at Stanford University to reminisce about Shockley's contributions to the information technology age. They unanimously concluded that Shockley was the man who brought silicon to Silicon Valley!

EXERCISE

7.15 A BJT having $\beta = 100$ is biased at a dc collector current of 1 mA. Find the value of g_m, r_e, and r_π at the bias point.
Ans. 40 mA/V; 25 Ω; 2.5 kΩ

The Voltage Gain The total collector voltage v_{CE} is

$$\begin{aligned} v_{CE} &= V_{CC} - i_C R_C \\ &= V_{CC} - (I_C + i_c) R_C \\ &= (V_{CC} - I_C R_C) - i_c R_C \\ &= V_{CE} - i_c R_C \end{aligned} \tag{7.76}$$

Thus, superimposed on the collector bias voltage V_{CE} we have signal voltage v_{ce} given by

$$\begin{aligned} v_{ce} &= -i_c R_C = -g_m v_{be} R_C \\ &= (-g_m R_C) v_{be} \end{aligned} \tag{7.77}$$

from which we find the voltage gain A_v of this amplifier as

$$A_v \equiv \frac{v_{ce}}{v_{be}} = -g_m R_C \tag{7.78}$$

Here again we note that because g_m is directly proportional to the collector bias current, the gain will be as stable as the collector bias current is made. Substituting for g_m from Eq. (7.62)

enables us to express the gain in the form

$$A_v = -\frac{I_C R_C}{V_T} \tag{7.79}$$

which is identical to the expression we derived in Section 7.1 (Eq. 7.21). Finally, we note that the gain expression in Eq. (7.78) is identical in form to that for the MOSFET amplifier (namely, $-g_m R_D$).

EXERCISE

7.16 In the circuit of Fig. 7.20(a), V_{BE} is adjusted to yield a dc collector current of 1 mA. Let $V_{CC} = 15$ V, $R_C = 10$ kΩ, and $\beta = 100$. Find the voltage gain v_{ce}/v_{be}. If $v_{be} = 0.005 \sin \omega t$ volt, find $v_C(t)$ and $i_B(t)$.
Ans. -400 V/V; $5 - 2 \sin \omega t$ volts; $10 + 2 \sin \omega t$ μA

Separating the Signal and the DC Quantities The analysis above indicates that every current and voltage in the amplifier circuit of Fig. 7.20(a) is composed of two components: a dc component and a signal component. For instance, $v_{BE} = V_{BE} + v_{be}$, $I_C = I_C + i_c$, and so on. The dc components are determined from the dc circuit given in Fig. 7.20(b) and from the relationships imposed by the transistor (Eqs. 7.52 through 7.54). On the other hand, a representation of the signal operation of the BJT can be obtained by eliminating the dc sources, as shown in Fig. 7.23. Observe that since the voltage of an ideal dc supply does not change, the signal voltage across it will be zero. For this reason we have replaced V_{CC} and V_{BE} with short circuits. Had the circuit contained ideal dc current sources, these would have been replaced by open circuits. Note, however, that the circuit of Fig. 7.23 is useful only insofar as it shows the various signal currents and voltages; it is *not* an actual amplifier circuit, since the dc bias circuit is not shown.

Figure 7.23 also shows the expressions for the current increments (i_c, i_b, and i_e) obtained when a small signal v_{be} is applied. These relationships can be represented by a circuit. Such

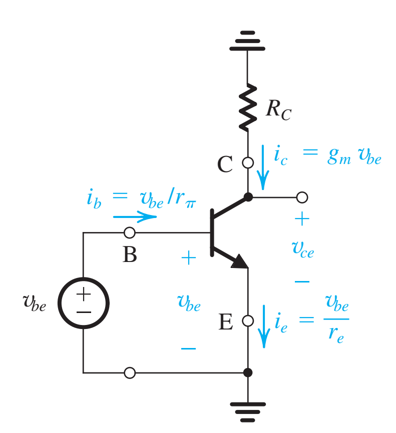


Figure 7.23 The amplifier circuit of Fig. 7.20(a) with the dc sources (V_{BE} and V_{CC}) eliminated (short-circuited). Thus only the signal components are present. Note that this is a representation of the signal operation of the BJT and not an actual amplifier circuit.

a circuit should have three terminals—C, B, and E—and should yield the same terminal currents indicated in Fig. 7.23. The resulting circuit is then *equivalent to the transistor as far as small-signal operation is concerned,* and thus it can be considered an equivalent small-signal circuit model.

The Hybrid-π Model An equivalent-circuit model for the BJT is shown in Fig. 7.24(a). This model represents the BJT as a voltage-controlled current source and explicitly includes the input resistance looking into the base, r_π. The model obviously yields $i_c = g_m v_{be}$ and $i_b = v_{be}/r_\pi$. Not so obvious, however, is the fact that the model also yields the correct expression for i_e. This can be shown as follows: At the emitter node we have

$$i_e = \frac{v_{be}}{r_\pi} + g_m v_{be} = \frac{v_{be}}{r_\pi}(1 + g_m r_\pi)$$

$$= \frac{v_{be}}{r_\pi}(1+\beta) = v_{be} \Big/ \left(\frac{r_\pi}{1+\beta}\right)$$

$$= v_{be}/r_e$$

A slightly different equivalent-circuit model can be obtained by expressing the current of the controlled source ($g_m v_{be}$) in terms of the base current i_b as follows:

$$g_m v_{be} = g_m (i_b r_\pi)$$

$$= (g_m r_\pi) i_b = \beta i_b$$

This results in the alternative equivalent-circuit model shown in Fig. 7.24(b). Here the transistor is represented as a current-controlled current source, with the control current being i_b.

As we have done in the case of the MOSFET's small-signal models, we can account for the Early effect (the slight dependence of i_C on v_{CE} due to basewidth modulation) by adding the resistance $r_o = V_A/I_C$ between collector and emitter, as shown in Fig. 7.25. Note that to conform with the literature, we have renamed v_{be} as v_π. The two models of Fig. 7.25 are versions of the hybrid-π model, the most widely used model for the BJT. The equivalent circuit of Fig. 7.25(a) corresponds to that of the MOSFET [Fig. 7.13(b)] except for r_π, which accounts for the finite base current (or finite β) of the BJT. However, the equivalent circuit of Fig. 7.25(b) has no MOS counterpart.

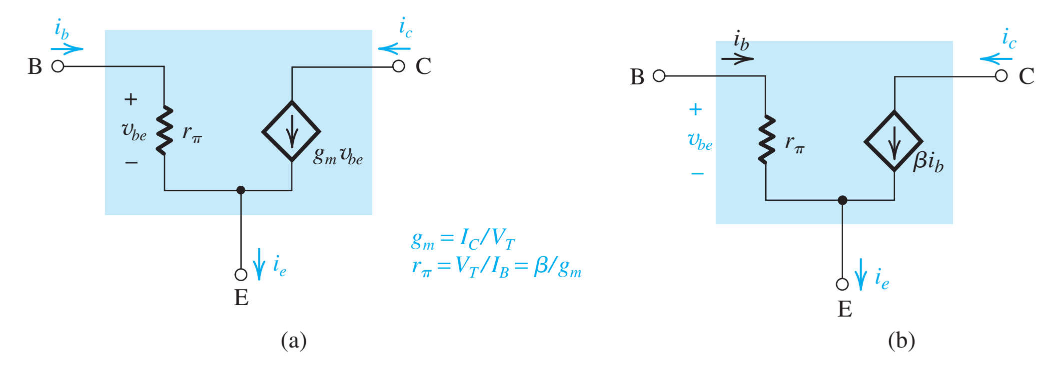


Figure 7.24 Two slightly different versions of the hybrid-π model for the small-signal operation of the BJT. The equivalent circuit in (**a**) represents the BJT as a voltage-controlled current source (a transconductance amplifier), and that in (**b**) represents the BJT as a current-controlled current source (a current amplifier).

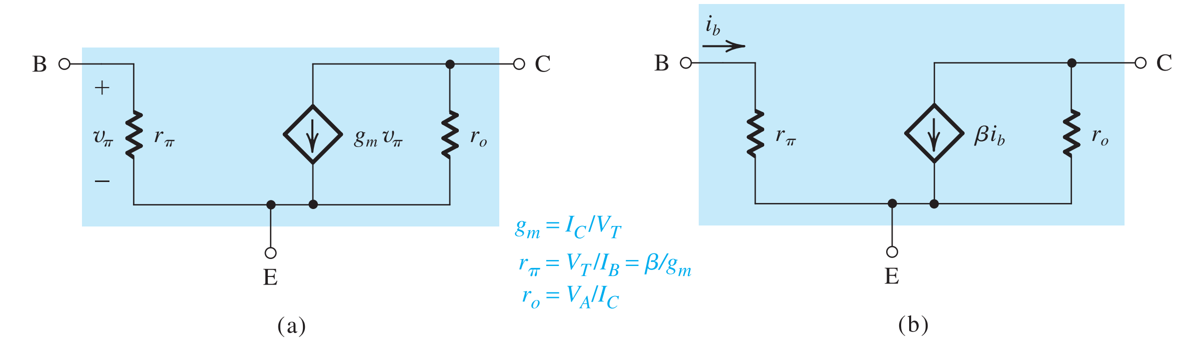


Figure 7.25 The hybrid-π small-signal model, in its two versions, with the resistance r_o included.

It is important to note that the small-signal equivalent circuits of Fig. 7.25 model the operation of the BJT *at a given bias point*. This should be obvious from the fact that the model parameters g_m, r_π, and r_o depend on the value of the dc bias current I_C, as indicated in Fig. 7.25. That is, these equivalent circuits model the *incremental operation* of the BJT around the bias point.

As in the case of the MOSFET amplifier, including r_o in the BJT model causes the voltage gain of the conceptual amplifier of Fig. 7.20(a) to become

$$\frac{v_o}{v_{be}} = -g_m(R_C \parallel r_o) \tag{7.80}$$

Thus, the magnitude of the gain is reduced somewhat.

EXERCISE

7.17 For the model in Fig. 7.24(b) show that $i_c = g_m v_{be}$ and $i_e = v_{be}/r_e$.

The T Model Although the hybrid-π model (in one of its two variants shown in Fig. 7.24) can be used to carry out small-signal analysis of any transistor circuit, there are situations in which an alternative model, shown in Fig. 7.26, is much more convenient. This model, called, as in the case of the MOSFET, the **T model**, is shown in two versions in Fig. 7.26. The model of Fig. 7.26(a) represents the BJT as a voltage-controlled current source with the control voltage being v_{be}. Here, however, the resistance between base and emitter, looking into the emitter, is explicitly shown. From Fig. 7.26(a) we see clearly that the model yields the correct expressions for i_c and i_e. It can also be shown to yield the correct expression for i_b (see Exercise 7.18 on the next page).

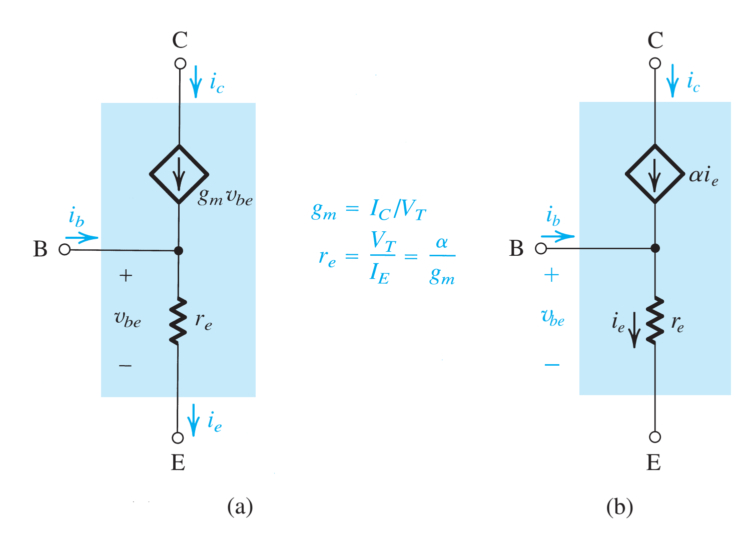


Figure 7.26 Two slightly different versions of what is known as the *T model* of the BJT. The circuit in (**a**) is a voltage-controlled current source representation and that in (**b**) is a current-controlled current source representation. These models explicitly show the emitter resistance r_e rather than the base resistance r_π featured in the hybrid-π model.

If in the model of Fig. 7.26(a) the current of the controlled source is expressed in terms of the emitter current as

$$\begin{aligned} g_m v_{be} &= g_m(i_e r_e) \\ &= (g_m r_e) i_e = \alpha i_e \end{aligned}$$

we obtain the alternative T model shown in Fig. 7.26(b). Here the BJT is represented as a current-controlled current source but with the control signal being i_e.

Finally, the T models can be augmented by r_o to account for the dependence of i_c to v_{ce} (the Early effect) to obtain the equivalent circuits shown in Fig. 7.27.

EXERCISE

7.18 Show that for the T model in Fig. 7.24(a), $i_b = v_{be}/r_\pi$.

Small-Signal Models of the *pnp* Transistor Although the small-signal models in Figs. 7.25 and 7.27 were developed for the case of the *npn* transistor, they apply equally well to the *pnp* transistor *with no change in polarities*.

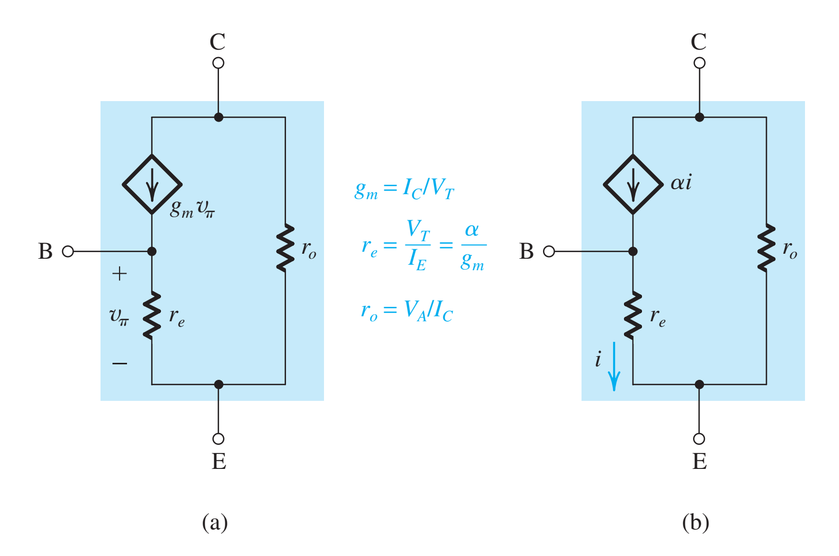


Figure 7.27 The T models of the BJT.

Example 7.5

We wish to analyze the transistor amplifier shown in Fig. 7.28(a) to determine its voltage gain v_o/v_i. Assume $\beta = 100$ and neglect the Early effect.

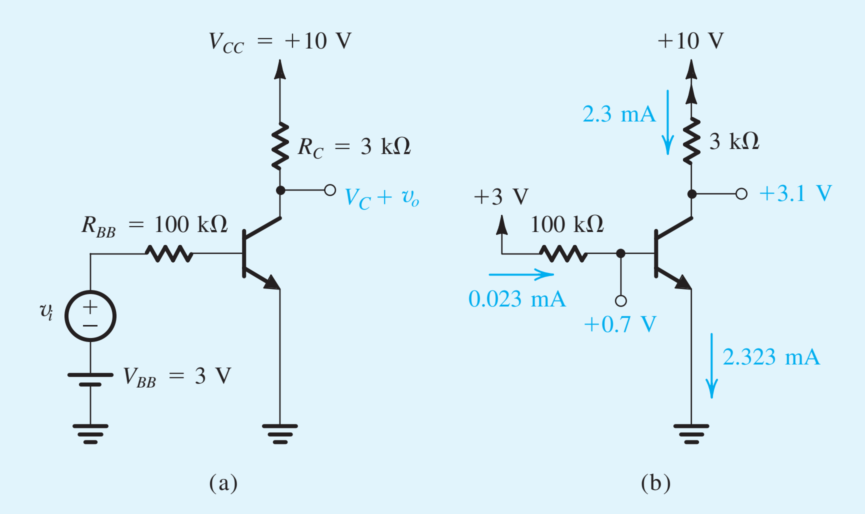


Figure 7.28 Example 7.5: (**a**) amplifier circuit; (**b**) circuit for dc analysis; (**c**) amplifier circuit with dc sources replaced by short circuits; (**d**) amplifier circuit with transistor replaced by its hybrid-π, small-signal model.

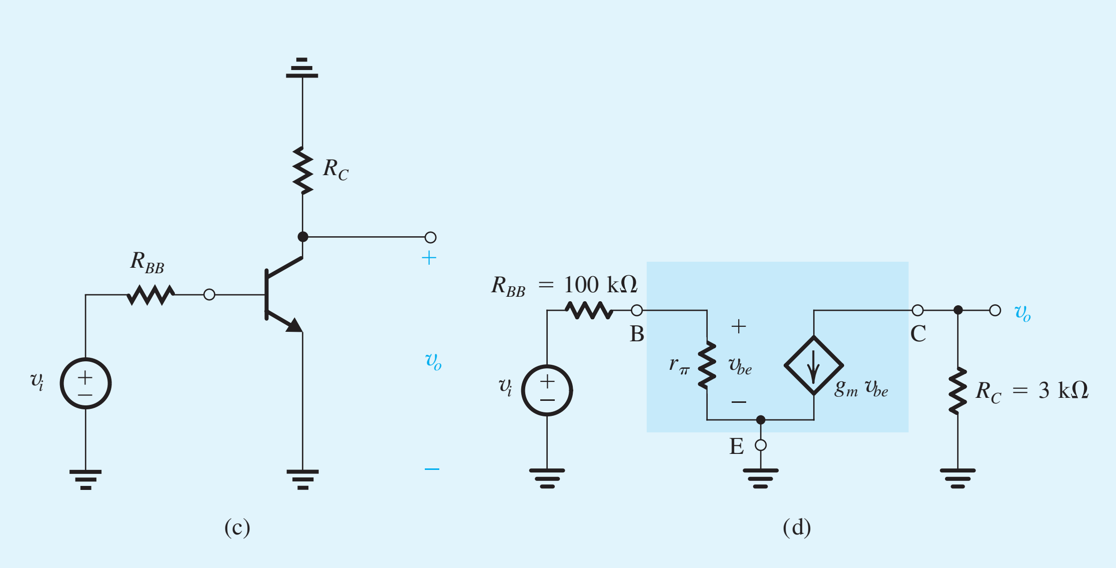


Figure 7.28 *continued*

Solution

We shall follow a five-step process:

1. The first step in the analysis consists of determining the quiescent operating point. For this purpose we assume that $v_i = 0$ and thus obtain the dc circuit shown in Fig. 7.28(b). The dc base current will be

$$I_B = \frac{V_{BB} - V_{BE}}{R_{BB}}$$
$$\simeq \frac{3 - 0.7}{100} = 0.023 \text{ mA}$$

The dc collector current will be

$$I_C = \beta I_B = 100 \times 0.023 = 2.3 \text{ mA}$$

The dc voltage at the collector will be

$$V_C = V_{CC} - I_C R_C$$
$$= +10 - 2.3 \times 3 = +3.1 \text{ V}$$

Since V_B at +0.7 V is less than V_C, it follows that in the quiescent condition the transistor will be operating in the active mode. The dc analysis is illustrated in Fig. 7.28(b).

Example 7.5 *continued*

2. Having determined the operating point, we can now proceed to determine the small-signal model parameters:

$$r_e = \frac{V_T}{I_E} = \frac{25 \text{ mV}}{(2.3/0.99) \text{ mA}} = 10.8\ \Omega$$
$$g_m = \frac{I_C}{V_T} = \frac{2.3 \text{ mA}}{25 \text{ mV}} = 92 \text{ mA/V}$$
$$r_\pi = \frac{\beta}{g_m} = \frac{100}{92} = 1.09 \text{ k}\Omega$$

3. Replacing V_{BB} and V_{CC} with short circuits results in the circuit in Fig. 7.28(c).
4. To carry out the small-signal analysis it is equally convenient to employ either of the two hybrid-π, equivalent-circuit models of Fig. 7.24 to replace the transistor in the circuit of Fig. 7.28(c). Using the first results in the amplifier equivalent circuit given in Fig. 7.28(d).
5. Analysis of the equivalent circuit in Fig. 7.28(d) proceeds as follows:

$$v_{be} = v_i \frac{r_\pi}{r_\pi + R_{BB}}$$
$$= v_i \frac{1.09}{101.09} = 0.011 v_i \tag{7.81}$$

The output voltage v_o is given by

$$v_o = -g_m v_{be} R_C$$
$$= -92 \times 0.011 v_i \times 3 = -3.04 v_i$$

Thus the voltage gain will be

$$A_v = \frac{v_o}{v_i} = -3.04 \text{ V/V} \tag{7.82}$$

Example 7.6

To gain more insight into the operation of transistor amplifiers, we wish to consider the waveforms at various points in the circuit analyzed in the previous example. For this purpose assume that v_i has a triangular waveform. First determine the maximum amplitude that v_i is allowed to have. Then, with the amplitude of v_i set to this value, give the waveforms of the total quantities $i_B(t)$, $v_{BE}(t)$, $i_C(t)$, and $v_C(t)$.

Solution

One constraint on signal amplitude is the small-signal approximation, which stipulates that v_{be} should not exceed about 10 mV. If we take the triangular waveform v_{be} to be 20 mV peak-to-peak and work backward, Eq. (7.81) can be used to determine the maximum possible peak of v_i,

$$\hat{v}_i = \frac{\hat{v}_{be}}{0.011} = \frac{10}{0.011} = 0.91 \text{ V}$$

To check whether the transistor remains in the active mode with v_i having a peak value $\hat{v}_i = 0.91$V, we have to evaluate the collector voltage. The voltage at the collector will consist of a triangular wave v_o superimposed on the dc value $V_C = 3.1$ V. The peak voltage of the triangular waveform will be

$$\hat{v}_o = \hat{v}_i \times \text{gain} = 0.91 \times 3.04 = 2.77 \text{ V}$$

It follows that when the output swings negative, the collector voltage reaches a minimum of $3.1 - 2.77 = 0.33$ V, which is lower than the base voltage by less than 0.4 V. Thus the transistor will remain in the active mode with v_i having a peak value of 0.91 V. Nevertheless, to be on the safe side, we will use a somewhat lower value for $\hat{v}_i$ of approximately 0.8 V, as shown in Fig. 7.29(a), and complete the analysis of this problem utilizing the equivalent circuit in Fig. 7.28(d). The signal current in the base will be triangular, with a peak value $\hat{i}_b$ of

$$\hat{i}_b = \frac{\hat{v}_i}{R_{BB} + r_\pi} = \frac{0.8}{100 + 1.09} = 0.008 \text{ mA}$$

This triangular-wave current will be superimposed on the quiescent base current I_B, as shown in Fig. 7.29(b). The base–emitter voltage will consist of a triangular-wave component superimposed on the dc V_{BE} that is approximately 0.7 V. The peak value of the triangular waveform will be

$$\hat{v}_{be} = \hat{v}_i \frac{r_\pi}{r_\pi + R_{BB}} = 0.8 \frac{1.09}{100 + 1.09} = 8.6 \text{ mV}$$

The total v_{BE} is sketched in Fig. 7.29(c).

The signal current in the collector will be triangular in waveform, with a peak value $\hat{i}_c$ given by

$$\hat{i}_c = \beta \hat{i}_b = 100 \times 0.008 = 0.8 \text{ mA}$$

This current will be superimposed on the quiescent collector current I_C $(= 2.3 \text{ mA})$, as shown in Fig. 7.29(d).

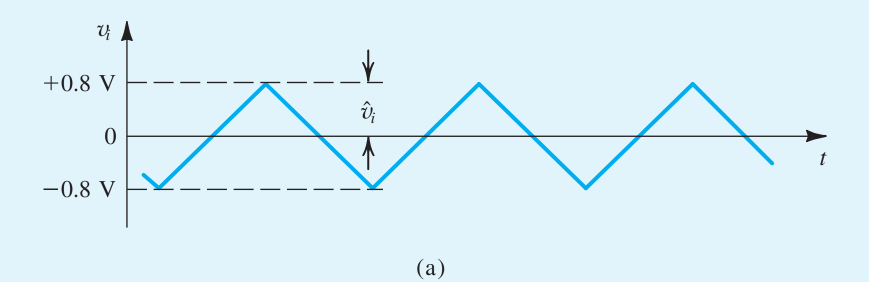
The signal voltage at the collector can be obtained by multiplying v_i by the voltage gain; that is,

$$\hat{v}_o = 3.04 \times 0.8 = 2.43 \text{ V}$$

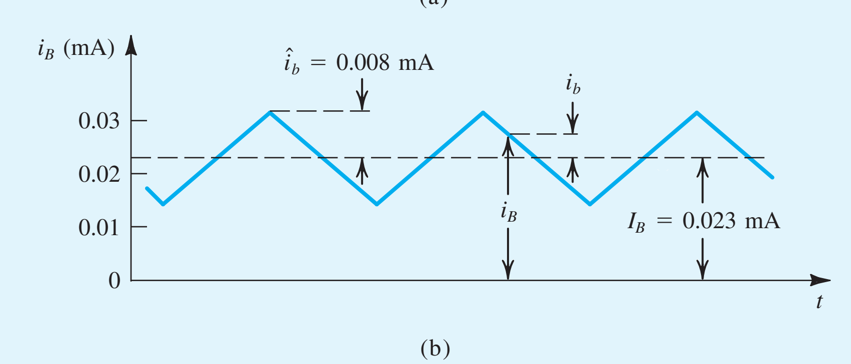
Figure 7.29(e) shows a sketch of the total collector voltage v_C versus time. Note the phase reversal between the input signal v_i and the output signal v_o.

Finally, we observe that each of the total quantities is the sum of a dc quantity (found from the dc circuit in Fig. 7.28b), and a signal quantity (found from the circuit in Fig. 7.28d).

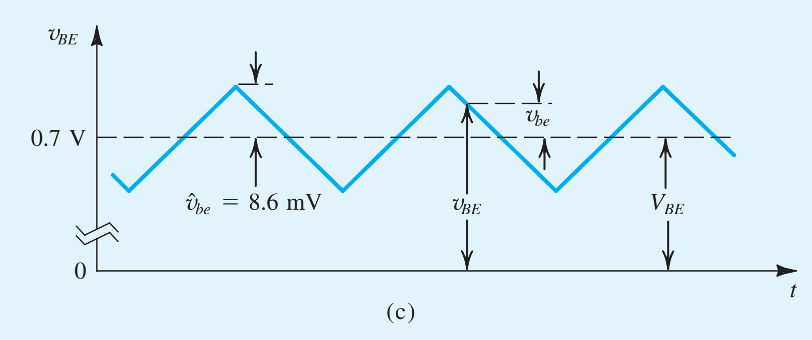
Example 7.6 *continued*



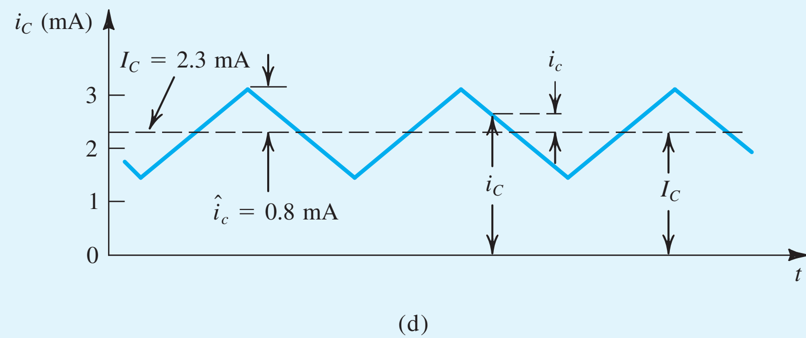
(a)



(b)



(c)



(d)

Figure 7.29 Signal waveforms in the circuit of Fig. 7.28.

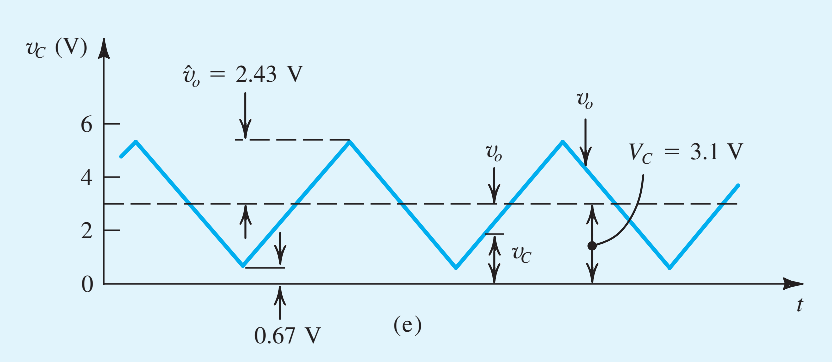


Figure 7.29 *continued*

Example 7.7

We need to analyze the circuit of Fig. 7.30(a) to determine the voltage gain and the signal waveforms at various points. The capacitor C_{C1} is a coupling capacitor whose purpose is to couple the signal v_i to the emitter while blocking dc. In this way the dc bias established by V^+ and V^- together with R_E and R_C will

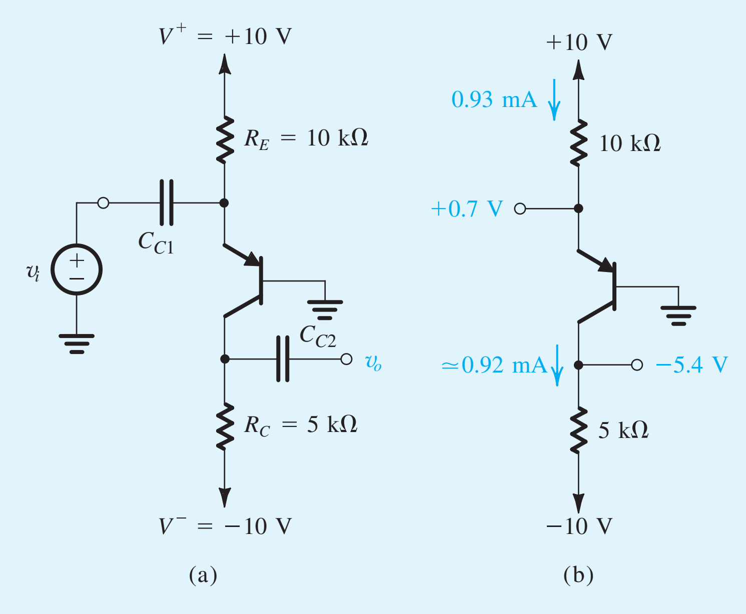


Figure 7.30 Example 7.7: (**a**) circuit; (**b**) dc analysis; (**c**) circuit with the dc sources eliminated; (**d**) small-signal analysis using the T model for the BJT.

Example 7.7 *continued*

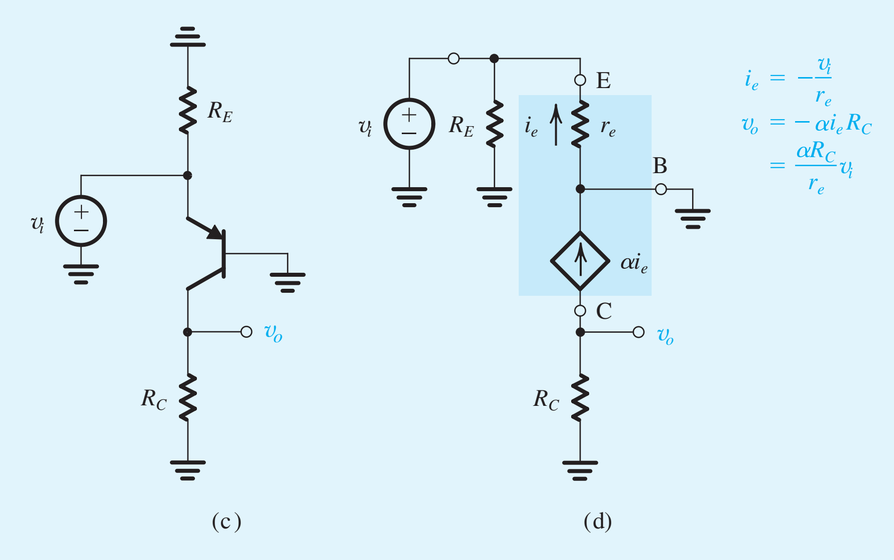


Figure 7.30 *continued*

not be disturbed when the signal v_i is connected. For the purpose of this example, C_{C1} will be assumed to be very large so as to act as a perfect short circuit at signal frequencies of interest. Similarly, another very large capacitor C_{C2} is used to couple the output signal v_o to other parts of the system. You may neglect the Early effect.

Solution

Here again we shall follow a five-step process:

1. Figure 7.30(b) shows the circuit with the signal source and the coupling capacitors eliminated. The dc operating point can be determined as follows:

$$I_E = \frac{+10 - V_E}{R_E} \simeq \frac{+10 - 0.7}{10} = 0.93 \text{ mA}$$

Assuming $\beta = 100$, then $\alpha = 0.99$, and

$$I_C = 0.99 I_E = 0.92 \text{ mA}$$
$$V_C = -10 + I_C R_C$$
$$= -10 + 0.92 \times 5 = -5.4 \text{ V}$$

Thus the transistor is in the active mode.

2. We now determine the small-signal parameters as follows:

$$g_m = \frac{I_C}{V_T} = \frac{0.92}{0.025} = 36.8 \text{ mA/V}$$

$$r_e = \frac{V_T}{I_E} = \frac{0.025}{0.92} = 27.2\ \Omega$$

$$\beta = 100 \qquad \alpha = 0.99$$

$$r_\pi = \frac{\beta}{g_m} = \frac{100}{36.8} = 2.72 \text{ k}\Omega$$

3. To prepare the circuit for small-signal analysis, we replace the dc sources with short circuits. The resulting circuit is shown in Fig. 7.30(c). Observe that we have also eliminated the two coupling capacitors, since they are assumed to be acting as perfect short circuits.
4. We are now ready to replace the BJT with one of the four equivalent-circuit models of Figs. 7.24 and 7.26. Although any of the four will work, the T models of Fig. 7.26 will be more convenient because the base is grounded. Selecting the version in Fig. 7.26(b) results in the amplifier equivalent circuit shown in Fig. 7.30(d).
5. Analysis of the circuit in Fig. 7.30(d) to determine the output voltage v_o and hence the voltage gain v_o/v_i is straightforward and is given in the figure. The result is

$$A_v = \frac{v_o}{v_i} = \frac{\alpha R_C}{r_e} = \frac{0.99 \times 5}{0.0272} = 182 \text{ V/V}$$

Note that the voltage gain is positive, indicating that the output is in phase with the input signal. This property is due to the fact that the input signal is applied to the emitter rather than to the base, as was done in Example 7.5. We should emphasize that the positive gain has nothing to do with the fact that the transistor used in this example is of the *pnp* type.

Returning to the question of allowable signal magnitude, we observe from Fig. 7.30(d) that $v_{eb} = v_i$. Thus, if small-signal operation is desired (for linearity), then the peak of v_i should be limited to approximately 10 mV. With $\hat{V}_i$ set to this value, as shown for a sine-wave input in Fig. 7.31, the peak amplitude at the collector, $\hat{V}_o$, will be

$$\hat{V}_o = 182 \times 0.01 = 1.82 \text{ V}$$

Example 7.7 *continued*

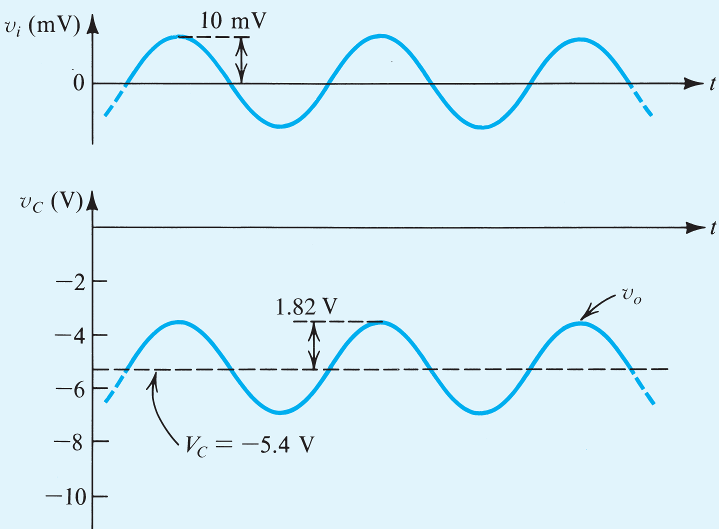


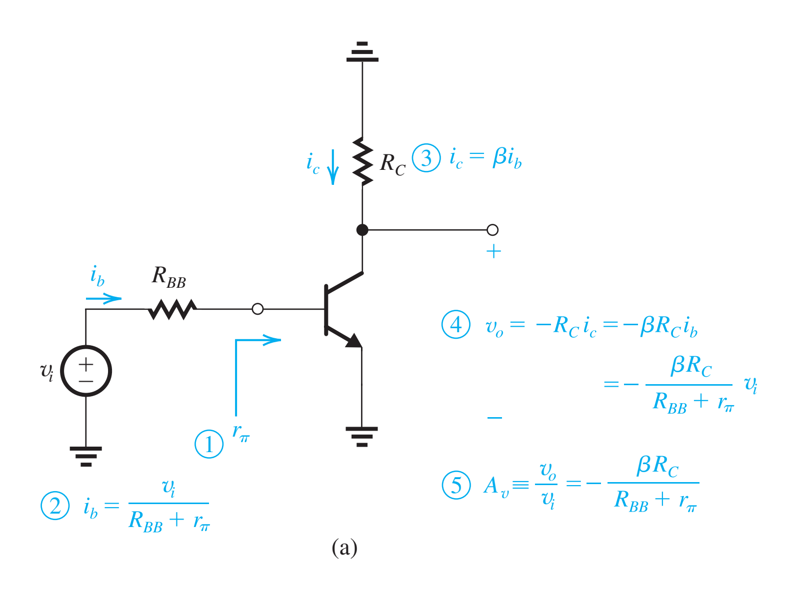
Figure 7.31 Input and output waveforms for the circuit of Fig. 7.30. Observe that this amplifier is noninverting, a property of the grounded-base configuration.

EXERCISE

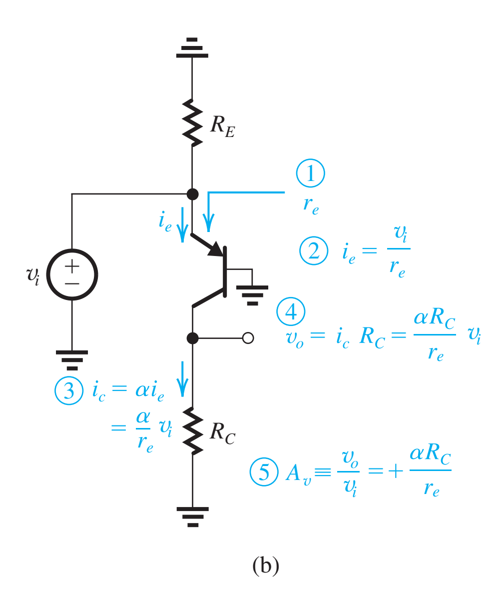
7.19 To increase the voltage gain of the amplifier analyzed in Example 7.7, the collector resistance R_C is increased to 7.5 kΩ. Find the new values of V_C, A_v, and the peak amplitude of the output sine wave corresponding to an input sine wave v_i of 10-mV peak.
Ans. –3.1 V; 276 V/V; 2.76 V

Performing Small-Signal Analysis Directly on the Circuit Diagram In most cases one should explicitly replace each BJT with its small-signal model and analyze the resulting circuit, as we have done in the examples above. This systematic procedure is particularly recommended for beginning students. Experienced circuit designers, however, often perform a first-order analysis directly on the circuit. Figure 7.32 illustrates this process for the two

circuits we analyzed in Examples 7.5 and 7.7. The reader is urged to follow this direct analysis procedure (the steps are numbered). Observe that the equivalent-circuit model is *implicitly* utilized; we are only saving the step of drawing the circuit with the BJT replaced by its model. Direct analysis, however, has an additional very important benefit: It provides insight regarding the signal transmission through the circuit. Such insight can prove invaluable in design, particularly at the stage of selecting a circuit configuration appropriate for a given application. Direct analysis can be utilized also for MOS amplifier circuits.



(a)



(b)

Figure 7.32 Performing signal analysis directly on the circuit diagram with the BJT small-signal model implicitly employed: (**a**) circuit for Example 7.5; (**b**) circuit for Example 7.7.

EXERCISE

7.20 The transistor in Fig. E7.20 is biased with a constant current source $I = 1$ mA and has $\beta = 100$ and $V_A = 100$ V.

(a) Neglecting the Early effect, find the dc voltages at the base, emitter, and collector.
(b) Find g_m, r_π, and r_o.
(c) If terminal Z is connected to ground, X to a signal source v_{sig} with a source resistance $R_{sig} =$ 2 kΩ, and Y to an 8-kΩ load resistance, use the hybrid-π model shown earlier (Fig. 7.25) to draw the small-signal equivalent circuit of the amplifier. (Note that the current source I should be replaced with an open circuit.) Calculate the overall voltage gain v_y/v_{sig}. If r_o is neglected, what is the error in estimating the gain magnitude? (*Note:* An infinite capacitance is used to indicate that the capacitance is sufficiently large that it acts as a short circuit at all signal frequencies of interest. However, the capacitor still blocks dc.)

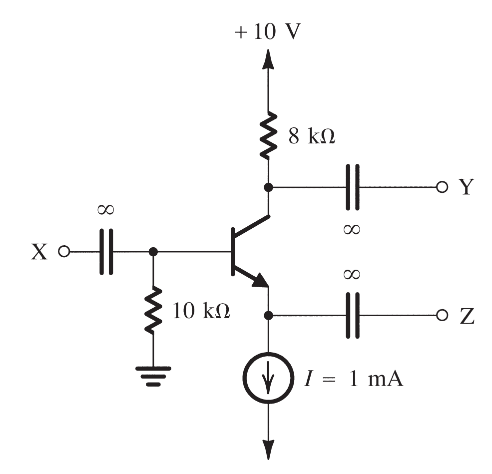


Figure E7.20

Ans. (a) –0.1 V, –0.8 V, +2.1 V; (b) 40 mA/V, 2.5 kΩ, 100 kΩ; (c) –77 V/V, +3.9%

7.2.3 Summary Tables

We conclude this section by presenting three useful summary tables: Table 7.1 lists the five steps to be followed in the analysis of a MOSFET or a BJT amplifier circuit. Table 7.2 presents the MOSFET small-signal, equivalent-circuit models, together with the formulas for calculating the parameter values of the models. Finally, Table 7.3 supplies the corresponding data for the BJT.

Table 7.1 Systematic Procedure for the Analysis of Transistor Amplifier Circuits

1. Eliminate the signal source and determine the dc operating point of the transistor.
2. Calculate the values of the parameters of the small-signal model.
3. Eliminate the dc sources by replacing each dc voltage source by a short circuit and each dc current source by an open circuit.
4. Replace the transistor with one of its small-signal, equivalent-circuit models. Although any of the models can be used, one might be more convenient than the others for the particular circuit being analyzed. This point will be made clearer in the next section.
5. Analyze the resulting circuit to determine the required quantities (e.g., voltage gain, input resistance).

Table 7.2 Small-Signal Models of the MOSFET

Small-Signal Parameters

NMOS transistors

- **Transconductance:**

$$g_m = \mu_n C_{ox}\frac{W}{L}V_{OV} = \sqrt{2\mu_n C_{ox}\frac{W}{L}I_D} = \frac{2I_D}{V_{OV}}$$

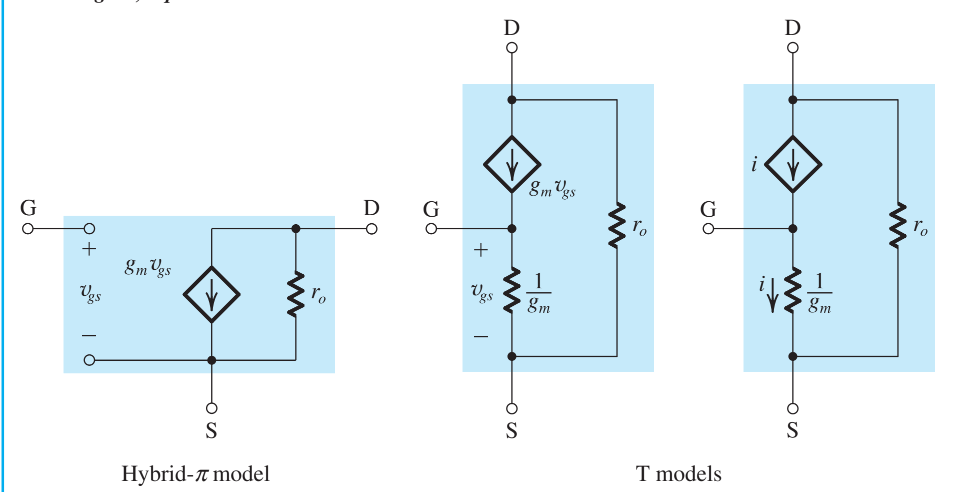
- Output resistance:

$$r_o = V_A/I_D = 1/\lambda I_D$$

PMOS transistors

Same formulas as for NMOS *except* using $|V_{OV}|$, $|V_A|$, $|\lambda|$ and replacing μ_n with μ_p.

Small-Signal, Equivalent-Circuit Models



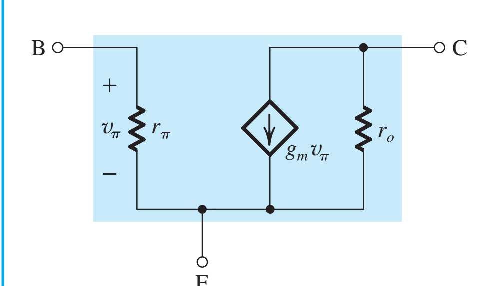
Hybrid-π model

T models

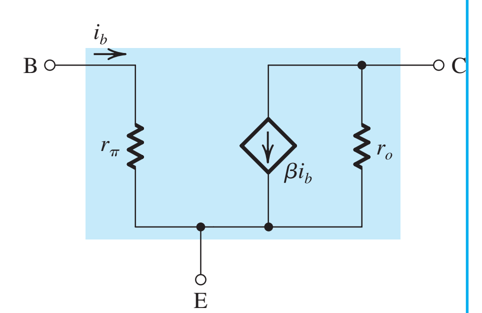
Table 7.3 Small-Signal Models of the BJT

Hybrid-π Model

- $(g_m v_\pi)$ Version

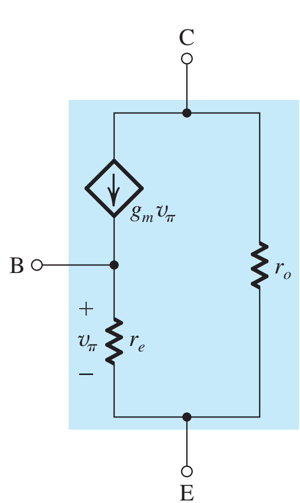


- (βi_b) Version

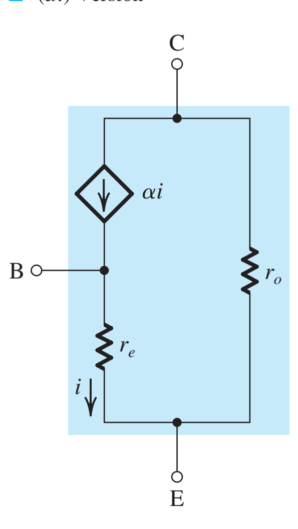


T Model

- $(g_m v_\pi)$ Version



- (αi) Version



Model Parameters in Terms of DC Bias Currents

$$g_m = \frac{I_C}{V_T} \qquad r_e = \frac{V_T}{I_E} = \alpha\frac{V_T}{I_C} \qquad r_\pi = \frac{V_T}{I_B} = \beta\frac{V_T}{I_C} \qquad r_o = \frac{|V_A|}{I_C}$$

In Terms of g_m

$$r_e = \frac{\alpha}{g_m} \qquad r_\pi = \frac{\beta}{g_m}$$

In Terms of r_e

$$g_m = \frac{\alpha}{r_e} \qquad r_\pi = (\beta + 1)r_e \qquad g_m + \frac{1}{r_\pi} = \frac{1}{r_e}$$

Relationships between α and β

$$\beta = \frac{\alpha}{1-\alpha} \qquad \alpha = \frac{\beta}{\beta+1} \qquad \beta + 1 = \frac{1}{1-\alpha}$$

7.3 Basic Configurations

It is useful at this point to take stock of where we are and where we are going in our study of transistor amplifiers. In Section 7.1 we examined the underlying principle for the application of the MOSFET, and of the BJT, as an amplifier. There we found that almost-linear amplification can be obtained by dc biasing the transistor at an appropriate point in its active region of operation, and by keeping the input signal (v_{gs} or v_{be}) small. We then developed, in Section 7.2, circuit models that represent the small-signal operation of each of the two transistor types (Tables 7.2 and 7.3), thus providing a systematic procedure (Table 7.1) for the analysis of transistor amplifiers.

We are now ready to consider the various possible configurations of MOSFET and BJT amplifiers, and we will do that in the present section. To focus our attention on the salient features of the various configurations, we shall present them in their most simple, or "stripped-down," version. Thus, we will not show the dc biasing arrangements, leaving the study of bias design to the next section. Finally, in Section 7.5 we will bring everything together and present practical *discrete-circuit amplifiers*, namely, amplifier circuits that can be constructed using discrete components. The study of integrated-circuit amplifiers begins in Chapter 8.

7.3.1 The Three Basic Configurations

There are three basic configurations for connecting a MOSFET or a BJT as an amplifier. Each of these configurations is obtained by connecting one of the device terminals to ground, thus creating a two-port network with the grounded terminal being *common* to the input and output ports. The resulting configurations are shown in Fig. 7.33(a–c) for the MOSFET and in Fig. 7.33(d–f) for the BJT.

In the circuit of Fig. 7.33(a) the source terminal is connected to ground, the input voltage signal v_i is applied between the gate and ground, and the output voltage signal v_o is taken between the drain and ground, across the resistance R_D. This configuration, therefore, is called the grounded-source or **common-source (CS)** amplifier. It is by far the most popular MOS amplifier configuration, and we utilized it in Sections 7.1 and 7.2 to study MOS amplifier operation. A parallel set of remarks apply to the BJT counterpart, the grounded-emitter or **common-emitter (CE)** amplifier in Fig. 7.33(d).

The **common-gate (CG)** or grounded-gate amplifier is shown in Fig. 7.33(b), and its BJT counterpart, the **common-base (CB)** or grounded-base amplifier in Fig. 7.33(e). Here the gate (base) is grounded, the input signal v_i is applied to the source (emitter), and the output signal v_o is taken at the drain (collector) across the resistance R_D (R_C). We encountered a CG amplifier in Example 7.4 and a CB amplifier in Example 7.7.

Finally, Fig. 7.33(c) shows the **common drain (CD)** or grounded-drain amplifier, and Fig. 7.31(f) shows its BJT counterpart, the **common-collector (CC)** or grounded collector amplifier. Here the drain (collector) terminal is grounded, the input signal v_i is applied between gate (base) and ground, and the output voltage v_o is taken between the source (emitter) and ground, across a resistance R_L. For reasons that will become apparent shortly, this pair of configurations is more commonly called the **source follower** and the **emitter follower**.

Our study of the three basic amplifier configurations of the MOSFET and of the BJT will reveal that each has distinctly different attributes, hence areas of application. As well, it will be shown that although each pair of configurations, (e.g., CS and CE), has many common attributes, important differences remain.

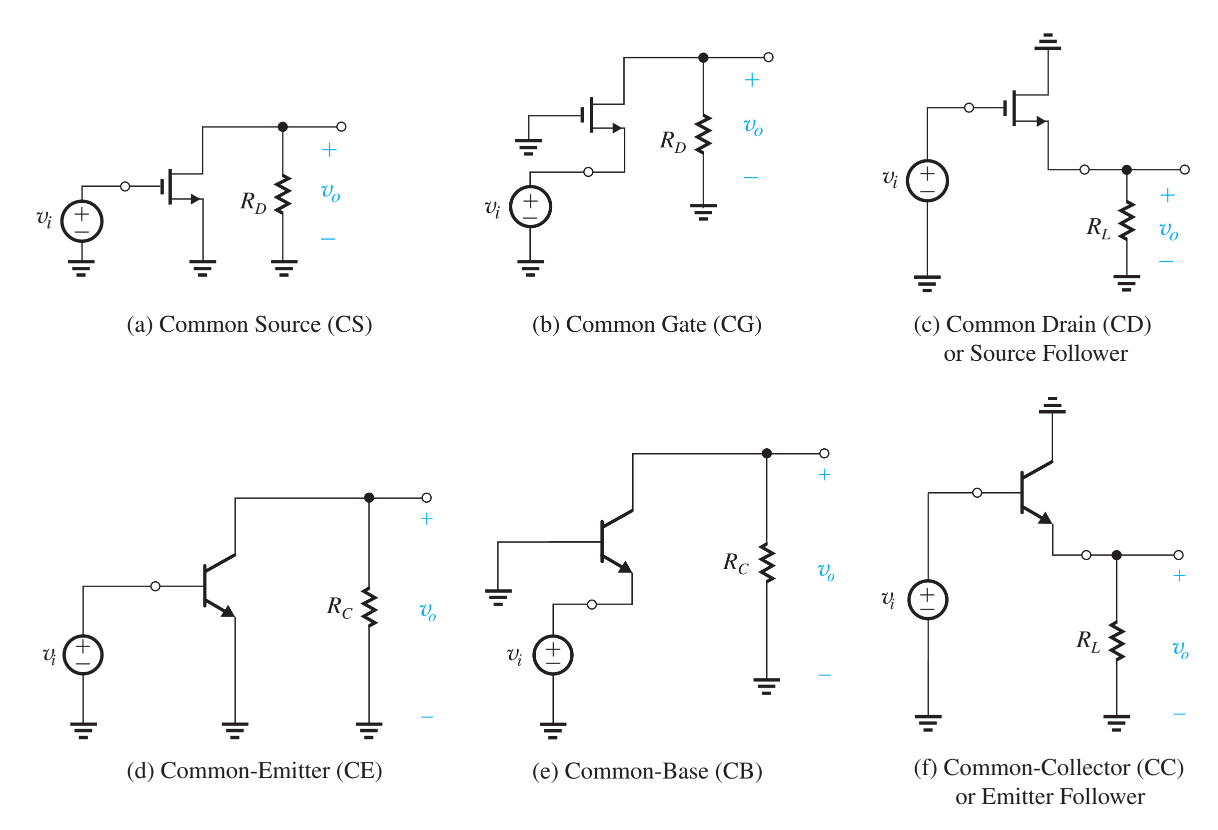


Figure 7.33 The basic configurations of transistor amplifiers. (**a**)–(**c**) For the MOSFET; (**d**)–(**f**) for the BJT.

Our next step is to replace the transistor in each of the six circuits in Fig. 7.33 by an appropriate equivalent-circuit model (from Tables 7.2 and 7.3) and analyze the resulting circuits to determine important characteristic parameters of the particular amplifier configuration. To simplify matters, we shall not include r_o in the initial analysis. At the end of the section we will offer a number of comments about when to include r_o in the analysis, and on the expected magnitude of its effect.

7.3.2 Characterizing Amplifiers

Before we begin our study of the different transistor amplifier configurations, we consider how to characterize the performance of an amplifier as a circuit building block. An introduction to this topic was presented in Section 1.5.

Figure 7.34(a) shows an amplifier fed with a signal source having an open-circuit voltage v_{sig} and an internal resistance R_{sig}. These can be the parameters of an actual signal source or, in a cascade amplifier, the Thévenin equivalent of the output circuit of another amplifier stage preceding the one under study. The amplifier is shown with a load resistance R_L connected to the output terminal. Here, R_L can be an actual load resistance or the input resistance of a succeeding amplifier stage in a cascade amplifier.

Figure 7.34(b) shows the amplifier circuit with the amplifier block replaced by its equivalent-circuit model. The input resistance R_{in} represents the loading effect of the amplifier

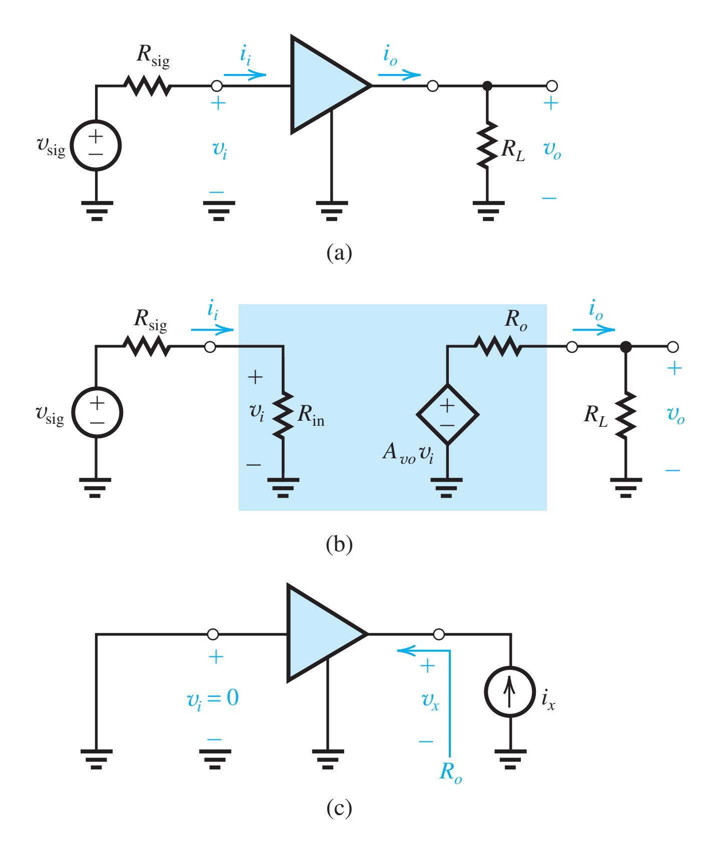


Figure 7.34 Characterization of the amplifier as a functional block: **(a)** An amplifier fed with a voltage signal v_{sig} having a source resistance R_{sig}, and feeding a load resistance R_L; **(b)** equivalent-circuit representation of the circuit in **(a)**; **(c)** determining the amplifier output resistance R_o.

input on the signal source. It is found from

$$R_{in} \equiv \frac{v_i}{i_i}$$

and together with the resistance R_{sig} forms a voltage divider that reduces v_{sig} to the value v_i that appears at the amplifier input,

$$v_i = \frac{R_{in}}{R_{in} + R_{sig}} v_{sig} \tag{7.83}$$

Most of the amplifier circuits studied in this section are **unilateral**. That is, they do not contain internal feedback, and thus R_{in} will be independent of R_L. However, in general R_{in} may depend on the load resistance R_L. Indeed one of the six configurations studied in this section, the emitter follower, exhibits such dependence.

The second parameter in characterizing amplifier performance is the **open-circuit voltage gain** A_{vo}, defined as

$$A_{vo} \equiv \left.\frac{v_o}{v_i}\right|_{R_L=\infty}$$

The third and final parameter is the output resistance R_o. Observe from Fig. 7.34(b) that R_o is the resistance seen looking back into the amplifier output terminal with v_i set to zero. Thus R_o can be determined, at least conceptually, as indicated in Fig. 7.34(c) with

$$R_o = \frac{v_x}{i_x}$$

Because R_o is determined with $v_i = 0$, the value of R_o does not depend on R_{sig}.

The controlled source $A_{vo}v_i$ and the output resistance R_o represent the Thévenin equivalent of the amplifier output circuit, and the output voltage v_o can be found from

$$v_o = \frac{R_L}{R_L + R_o} A_{vo} v_i \tag{7.84}$$

Thus the **voltage gain of the amplifier proper**, A_v, can be found as

$$A_v \equiv \frac{v_o}{v_i} = A_{vo} \frac{R_L}{R_L + R_o} \tag{7.85}$$

and the **overall voltage gain**, G_v,

$$G_v \equiv \frac{v_o}{v_{sig}}$$

can be determined by combining Eqs. (7.83) and (7.85):

$$G_v = \frac{R_{in}}{R_{in} + R_{sig}} A_{vo} \frac{R_L}{R_L + R_o} \tag{7.86}$$

7.3.3 The Common-Source (CS) and Common-Emitter (CE) Amplifiers

Of the three basic transistor amplifier configurations, the common-source (common-emitter, for BJT), is the most widely used. Typically, in an amplifier formed by cascading a number of gain stages, the bulk of the voltage gain is obtained by using one or more common-source (or common-emitter, for BJT) stages in cascade.

Characteristic Parameters of the CS Amplifier Figure 7.35(a) shows a common-source amplifier (with the biasing arrangement omitted) fed with a signal source v_{sig} having a source resistance R_{sig}. We wish to analyze this circuit to determine R_{in}, A_{vo}, and R_o. For this purpose, we assume that R_D is part of the amplifier; thus if a load resistance R_L is connected to the amplifier output, R_L appears in parallel with R_D. In such a case, we wish to determine A_v and G_v as well.

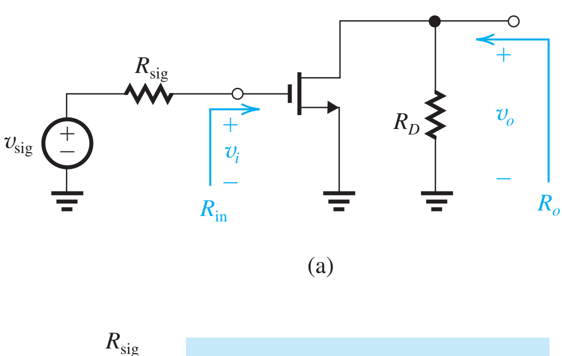
Replacing the MOSFET with its hybrid-π model (without r_o), we obtain the CS amplifier equivalent circuit in Fig. 7.35(b) for which, tracing the signal from input to output, we can write by inspection

$$R_{in} = \infty \tag{7.87}$$

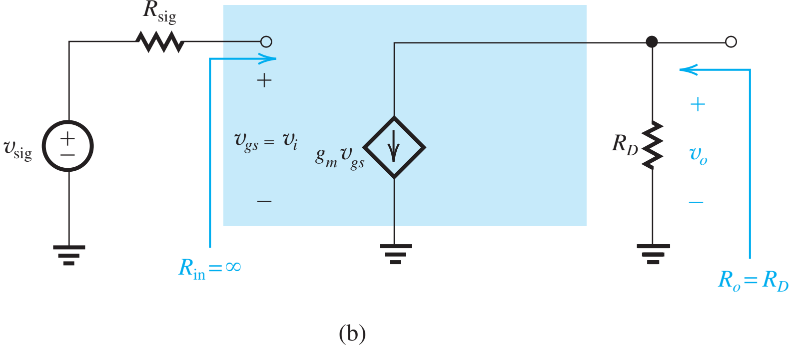
$$v_i = v_{sig}$$

$$v_{gs} = v_i$$

$$v_o = -g_m v_{gs} R_D$$



(a)



(b)

Figure 7.35 **(a)** Common-source amplifier fed with a signal v_{sig} from a generator with a resistance R_{sig}. The bias circuit is omitted. **(b)** The common-source amplifier with the MOSFET replaced with its hybrid-π model.

Thus,

$$A_{vo} \equiv \frac{v_o}{v_i} = -g_m R_D \tag{7.88}$$

$$R_o = R_D \tag{7.89}$$

If a load resistance R_L is connected across R_D, the voltage gain A_v can be obtained from

$$A_v = A_{vo} \frac{R_L}{R_L + R_o} \tag{7.90}$$

where A_{vo} is given by Eq. (7.88) and R_o by Eq. (7.89), or alternatively by simply adding R_L in parallel with R_D in Eq. (7.88), thus

$$A_v = -g_m (R_D \| R_L) \tag{7.91}$$

The reader can easily show that the expression obtained from Eq. (7.90) is identical to that in Eq. (7.91). Finally, since $R_{in} = \infty$ and thus $v_i = v_{sig}$, the overall voltage gain G_v is equal to A_v,

$$G_v \equiv \frac{v_o}{v_{sig}} = -g_m (R_D \| R_L) \tag{7.92}$$

EXERCISE

7.21 A CS amplifier utilizes a MOSFET biased at $I_D = 0.25$ mA with $V_{OV} = 0.25$ V and $R_D = 20$ kΩ. The amplifier is fed with a signal source having $R_{sig} = 100$ kΩ, and a 20-kΩ load is connected to the output. Find R_{in}, A_{vo}, R_o, A_v, and G_v. If, to maintain reasonable linearity, the peak of the input sine-wave signal is limited to 10% of $2V_{OV}$, what is the peak of the sine-wave voltage at the output?
Ans. ∞; −40 V/V; 20 kΩ; −20 V/V; −20 V/V; 1 V

Characteristic Parameters of the CE Amplifier Figure 7.36(a) shows a common-emitter amplifier. Its equivalent circuit, obtained by replacing the BJT with its hybrid-π model (without r_o), is shown in Fig. 7.36(b). The latter circuit can be analyzed to obtain the characteristic parameters of the CE amplifier. The analysis parallels that for the MOSFET above except that here we have the added complexity of a finite input resistance r_π. Tracing the signal through the amplifier from input to output, we can write by inspection

$$R_{in} = r_\pi$$

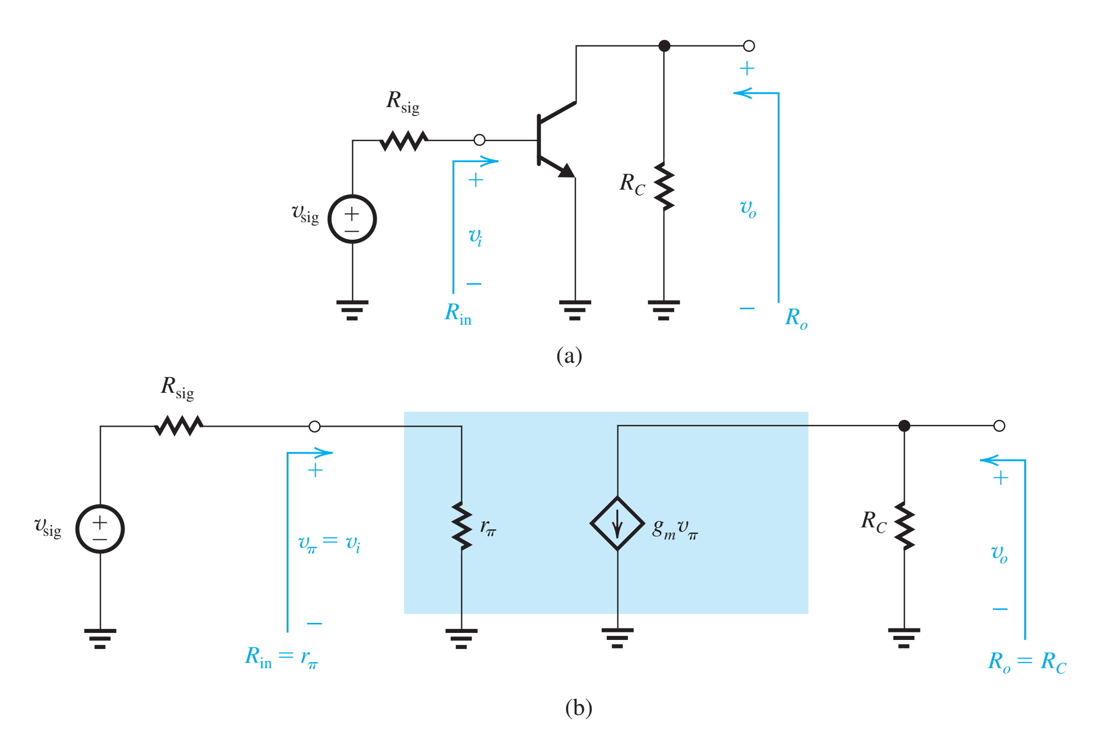


Figure 7.36 (**a**) Common-emitter amplifier fed with a signal v_{sig} from a generator with a resistance R_{sig}. The bias circuit is omitted. (**b**) The common-emitter amplifier circuit with the BJT replaced by its hybrid-π model.

Then we write

$$v_i = \frac{r_\pi}{r_\pi + R_{sig}} v_{sig} \tag{7.93}$$

$$v_\pi = v_i$$

$$v_o = -g_m v_\pi R_C$$

Thus,

$$A_{vo} \equiv \frac{v_o}{v_i} = -g_m R_C \tag{7.94}$$

$$R_o = R_C \tag{7.95}$$

With a load resistance R_L connected across R_C,

$$A_v = -g_m (R_C \| R_L) \tag{7.96}$$

and the overall voltage gain G_v can be found from

$$G_v \equiv \frac{v_o}{v_{sig}} = \frac{v_i}{v_{sig}} \frac{v_o}{v_i}$$

Thus,

$$G_v = -\frac{r_\pi}{r_\pi + R_{sig}} g_m (R_C \| R_L) \tag{7.97}$$

It is important to note here the effect of the finite input resistance (r_π) in reducing the magnitude of the voltage gain by the voltage-divider ratio $r_\pi/(r_\pi + R_{sig})$. The extent of the gain reduction depends on the relative values of r_π and R_{sig}. However, there is a compensating effect in the CE amplifier: g_m of the BJT is usually much higher than the corresponding value of the MOSFET.

Example 7.8

A CE amplifier utilizes a BJT with $\beta = 100$ is biased at $I_C = 1$ mA and has a collector resistance $R_C = 5$ kΩ. Find R_{in}, R_o, and A_{vo}. If the amplifier is fed with a signal source having a resistance of 5 kΩ, and a load resistance $R_L = 5$ kΩ is connected to the output terminal, find the resulting A_v and G_v. If $\hat{v}_\pi$ is to be limited to 5 mV, what are the corresponding $\hat{v}_{sig}$ and $\hat{v}_o$ with the load connected?

Solution

At $I_C = 1$ mA,

$$g_m = \frac{I_C}{V_T} = \frac{1 \text{ mA}}{0.025 \text{ V}} = 40 \text{ mA/V}$$

$$r_\pi = \frac{\beta}{g_m} = \frac{100}{40 \text{ mA/V}} = 2.5 \text{ k}\Omega$$

Example 7.8 *continued*

The amplifier characteristic parameters can now be found as

$$R_{\text{in}} = r_\pi = 2.5\ \text{k}\Omega$$
$$A_{vo} = -g_m R_C$$
$$= -40\ \text{mA/V} \times 5\ \text{k}\Omega$$
$$= -200\ \text{V/V}$$
$$R_o = R_C = 5\ \text{k}\Omega$$

With a load resistance $R_L = 5\ \text{k}\Omega$ connected at the output, we can find A_v by either of the following two approaches:

$$A_v = A_{vo}\frac{R_L}{R_L + R_o}$$
$$= -200 \times \frac{5}{5+5} = -100\ \text{V/V}$$

or

$$A_v = -g_m(R_C \,\|\, R_L)$$
$$= -40(5 \,\|\, 5) = -100\ \text{V/V}$$

The overall voltage gain G_v can now be determined as

$$G_v = \frac{R_{\text{in}}}{R_{\text{in}} + R_{\text{sig}}} A_v$$
$$= \frac{2.5}{2.5+5} \times -100 = -33.3\ \text{V/V}$$

If the maximum amplitude of v_π is to be 5 mV, the corresponding value of $\hat{v}_{\text{sig}}$ will be

$$\hat{v}_{\text{sig}} = \left(\frac{R_{\text{in}} + R_{\text{sig}}}{R_{\text{in}}}\right)\hat{v}_\pi = \frac{2.5+5}{2.5} \times 5 = 15\ \text{mV}$$

and the amplitude of the signal at the output will be

$$\hat{v}_o = G_v \hat{v}_{\text{sig}} = 33.3 \times 0.015 = 0.5\ \text{V}$$

EXERCISE

7.22 The designer of the amplifier in Example 7.8 decides to lower the bias current to half its original value in order to raise the input resistance and hence increase the fraction of v_{sig} that appears at the input of the amplifier proper. In an attempt to maintain the voltage gain, the designer decides to double the value of R_C. For the new design, determine R_{in}, A_{vo}, R_o, A_v, and G_v. If the peak amplitude of v_π is to be limited to 5 mV, what are the corresponding values of $\hat{v}_{sig}$ and $\hat{v}_o$ (with the load connected)?
Ans. 5 kΩ; −200 V/V; 10 kΩ; −66.7 V/V; −33.3 V/V; 10 mV; 0.33 V

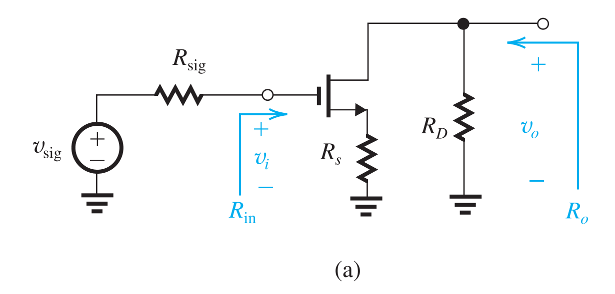
Comment: Although a larger fraction of the input signal reaches the amplifier input, linearity considerations cause the output signal to be in fact smaller than in the original design!

Final Remarks

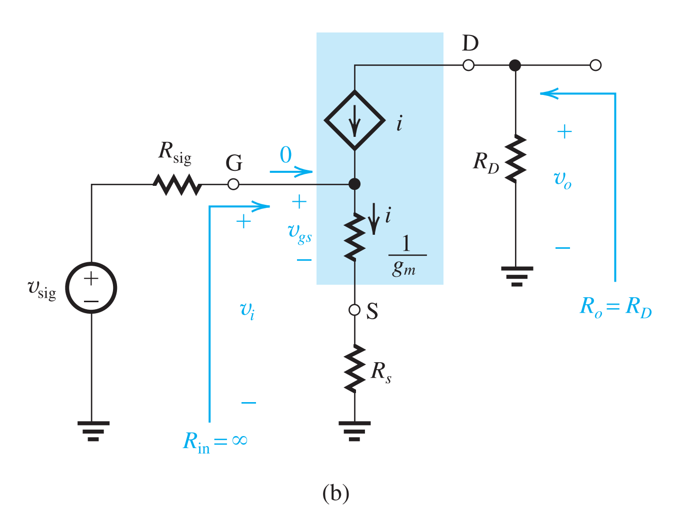
1. The CS and CE amplifiers are the most useful of all transistor amplifier configurations. They exhibit a moderate to high input resistance (infinite for the CS), a moderate to high output resistance, and reasonably high voltage gain.
2. The input resistance of the CE amplifier, $R_{in} = r_\pi = \beta/g_m$, is inversely proportional to the dc bias current I_C. To increase R_{in} one is tempted to lower the bias current I_C; however, this also lowers g_m and hence the voltage gain. This is a significant design trade-off. If a much higher input resistance is desired, then a modification of the CE configuration (to be discussed in Section 7.3.4) can be applied, or an emitter-follower stage can be inserted between the signal source and the CE amplifier (see Section 7.3.6).
3. Reducing R_D or R_C to lower the output resistance of the CS or CE amplifier, respectively, is usually not a viable proposition because the voltage gain is also reduced. Alternatively, if a very low output resistance (in the ohms or tens-of-ohms range) is needed, a source-follower or an emitter-follower stage can be utilized between the output of the CS or CE amplifier and the load resistance (see Section 7.3.6).
4. Although the CS and the CE configurations are the workhorses of transistor amplifiers, both suffer from a limitation on their high-frequency response. As will be shown in Chapter 10, combining the CS (CE) amplifier with a CG (CB) amplifier can extend the bandwidth considerably. The CG and CB amplifiers are studied in Section 7.3.5.

7.3.4 The Common-Source (Common-Emitter) Amplifier with a Source (Emitter) Resistance

It is often beneficial to insert a resistance R_s (a resistance R_e) in the source lead (the emitter lead) of a common-source (common-emitter) amplifier. Figure 7.37(a) shows a CS amplifier with a resistance R_s in its source lead. The corresponding small-signal equivalent circuit is shown



(a)



(b)

Figure 7.37 The CS amplifier with a source resistance R_s: **(a)** circuit without bias details; **(b)** equivalent circuit with the MOSFET represented by its T model.

in Fig. 7.37(b), where we have utilized the T model for the MOSFET. The T model is used in preference to the hybrid-π model because it makes the analysis in this case considerably simpler. In general, *whenever a resistance is connected in the source lead, the T model is preferred*. The source resistance then simply appears in series with the model resistance $1/g_m$ and can be added to it.

From Fig. 7.37(b) we see that as expected, the input resistance R_{in} is infinite and thus $v_i = v_{sig}$. Unlike the CS amplifier, however, here only a fraction of v_i appears between gate and source as v_{gs}. The voltage divider composed of $1/g_m$ and R_s, which appears across the amplifier input, can be used to determine v_{gs}, as follows:

$$v_{gs} = v_i \frac{1/g_m}{1/g_m + R_s} = \frac{v_i}{1 + g_m R_s} \tag{7.98}$$

Thus we can use the value of R_s to control the magnitude of the signal v_{gs} and thereby ensure that v_{gs} does not become too large and cause unacceptably high nonlinear distortion. This is the first benefit of including resistor R_s. Other benefits will be encountered in later sections and chapters. For instance, it will be shown in Chapter 10 that R_s causes the useful bandwidth of the amplifier to be extended. The mechanism by which R_s causes such improvements in amplifier performance is *negative feedback*. To see how R_s introduces negative feedback, refer to Fig. 7.37(a): If with v_{sig} and hence v_i kept constant, the drain current increases for some

reason, the source current also will increase, resulting in an increased voltage drop across R_s. Thus the source voltage rises, and the gate-to-source voltage decreases. The latter effect causes the drain current to decrease, counteracting the initially assumed change, an indication of the presence of negative feedback. In Chapter 11 we shall study negative feedback formally. There we will learn that the improvements that negative feedback provides are obtained at the expense of a reduction in gain. We will now show this to be the case in the circuit of Fig. 7.37.

The output voltage v_o is obtained by multiplying the controlled-source current i by R_D,

$$v_o = -iR_D$$

The current i in the source lead can be found by dividing v_i by the total resistance in the source,

$$i = \frac{v_i}{1/g_m + R_s} = \left(\frac{g_m}{1 + g_m R_s}\right) v_i \tag{7.99}$$

Thus, the voltage gain A_{vo} can be found as

$$A_{vo} \equiv \frac{v_o}{v_i} = -\frac{g_m R_D}{1 + g_m R_s} \tag{7.100}$$

which can also be expressed as

$$A_{vo} = -\frac{R_D}{1/g_m + R_s} \tag{7.101}$$

Equation (7.100) indicates that including the resistance R_s reduces the voltage gain by the factor $(1 + g_m R_s)$. This is the price paid for the improvements that accrue as a result of R_s. It is interesting to note that in Chapter 11, we will find that the factor $(1 + g_m R_s)$ is the "amount of negative feedback" introduced by R_s. It is also the same factor by which linearity, bandwidth, and other performance parameters improve. Because of the negative-feedback action of R_s it is known as a **source-degeneration resistance**.

There is another useful interpretation of the expression for the drain current in Eq. (7.99): The quantity between brackets on the right-hand side can be thought of as the "effective transconductance with R_s included." Thus, including R_s reduces the transconductance by the factor $(1 + g_m R_s)$. This, of course, is simply the result of the fact that only a fraction $1/(1 + g_m R_s)$ of v_i appears as v_{gs} (see Eq. 7.98).

The alternative gain expression in Eq. (7.101) has a powerful and insightful interpretation: The voltage gain between gate and drain is equal to the ratio of the total resistance in the drain (R_D) to the total resistance in the source ($1/g_m + R_s$),

$$\text{Voltage gain from gate to drain} = -\frac{\text{Total resistance in drain}}{\text{Total resistance in source}} \tag{7.102}$$

This is a general expression. For instance, setting $R_s = 0$ in Eq. (7.101) yields A_{vo} of the CS amplifier.

Finally, we consider the situation of a load resistance R_L connected at the output. We can obtain the gain A_v using the open-circuit voltage gain A_{vo} together with the output resistance R_o, which can be found by inspection to be

$$R_o = R_D$$

Alternatively, A_v can be obtained by simply replacing R_D in Eq. (7.101) or (7.100) by $(R_D \| R_L)$; thus,

$$A_v = -\frac{g_m(R_D \| R_L)}{1 + g_m R_s} \tag{7.103}$$

or

$$A_v = -\frac{R_D \| R_L}{1/g_m + R_s} \tag{7.104}$$

Observe that Eq. (7.104) is a direct application of the ratio of total resistance rule of Eq. (7.102). Finally, note that because R_{in} is infinite, $v_i = v_{sig}$ and the overall voltage gain G_v is equal to A_v.

EXERCISE

7.23 In Exercise 7.21 we applied an input signal v_{sig} of 50 mV peak and obtained an output signal of approximately 1 V peak. Assume that for some reason we now have an input signal v_{sig} that is 0.2 V peak and that we wish to modify the circuit to keep v_{gs} unchanged, and thus keep the nonlinear distortion from increasing. What value should we use for R_s? What value of G_v will result? What will the peak signal at the output become? Assume $r_o = \infty$.
Ans. 1.5 kΩ; −5 V/V; 1 V

We next turn our attention to the BJT case. Figure 7.38(a) shows a CE amplifier with a resistance R_e in its emitter. The corresponding equivalent circuit, utilizing the T model, is shown in Fig. 7.38(b). Note that in the BJT case also, as a general rule, the T model results in a simpler analysis and should be employed whenever there is a resistance in series with the emitter.

To determine the amplifier input resistance R_{in}, we note from Fig. 7.38(b) that

$$R_{in} \equiv \frac{v_i}{i_b}$$

where

$$i_b = (1-\alpha)i_e = \frac{i_e}{\beta + 1} \tag{7.105}$$

and

$$i_e = \frac{v_i}{r_e + R_e} \tag{7.106}$$

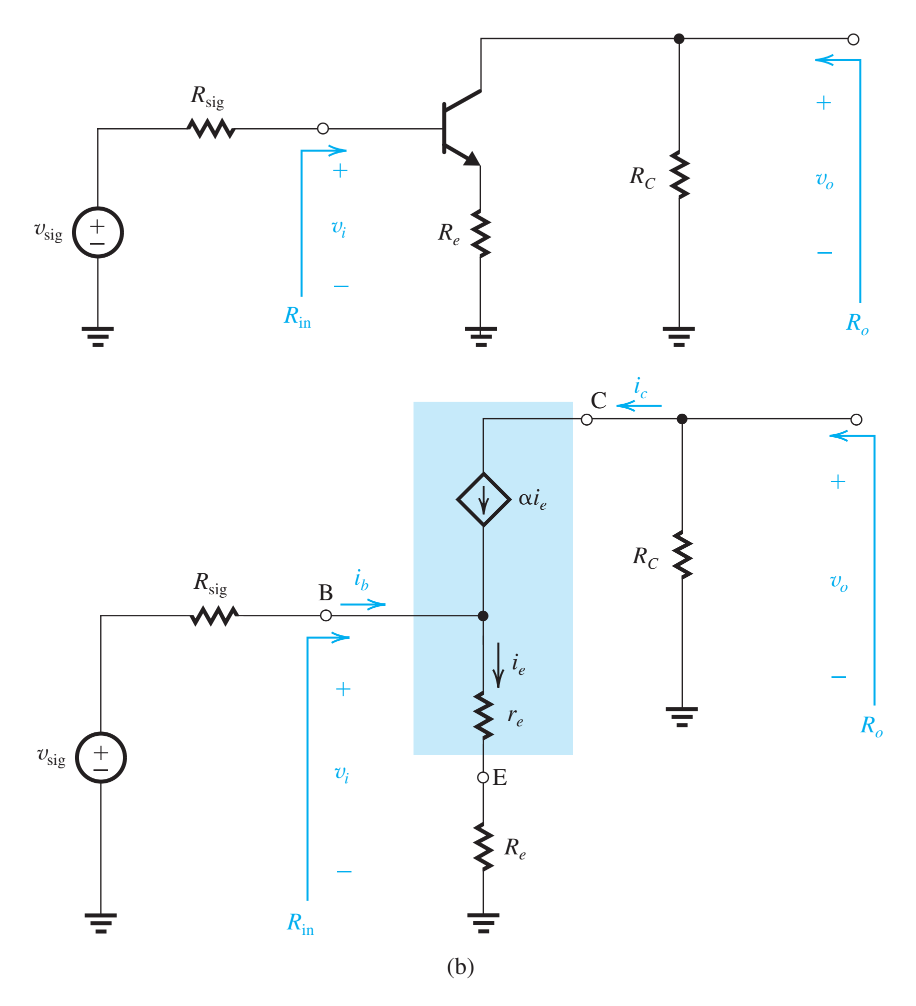


Figure 7.38 The CE amplifier with an emitter resistance R_e; (**a**) circuit without bias details; (**b**) equivalent circuit with the BJT replaced with its T model.

Thus,

$$R_{in} = (\beta + 1)(r_e + R_e) \tag{7.107}$$

This is a very important result. It states that *the input resistance looking into the base is* $(\beta + 1)$ *times the total resistance in the emitter*, and is known as the **resistance-reflection rule**. The factor $(\beta + 1)$ arises because the base current is $1/(\beta + 1)$ times the emitter current. The expression for R_{in} in Eq. (7.107) shows clearly that including a resistance R_e in the emitter can substantially increase R_{in}, a very desirable result. Indeed, the value of R_{in} is increased by the ratio

$$\frac{R_{in}(\text{with } R_e \text{ included})}{R_{in}(\text{without } R_e)} = \frac{(\beta + 1)(r_e + R_e)}{(\beta + 1)r_e}$$

$$= 1 + \frac{R_e}{r_e} \simeq 1 + g_m R_e \tag{7.108}$$

Thus the circuit designer can use the value of R_e to control the value of R_{in}.

To determine the voltage gain A_{vo}, we see from Fig. 7.38(b) that

$$v_o = -i_c R_C$$
$$= -\alpha i_e R_C$$

Substituting for i_e from Eq. (7.106) gives

$$A_{vo} = -\alpha \frac{R_C}{r_e + R_e} \tag{7.109}$$

This is a very useful result: It states that the gain from base to collector is α times the ratio of the total resistance in the collector to the total resistance in the emitter (in this case, $r_e + R_e$),

$$\text{Voltage gain from base to collector} = -\alpha \frac{\text{Total resistance in collector}}{\text{Total resistance in emitter}} \tag{7.110}$$

This is the BJT version of the MOSFET expression in Eq. (7.102) except that here we have the additional factor α. This factor arises because $i_c = \alpha i_e$, unlike the MOSFET case where $i_d = i_s$. Usually, $\alpha \simeq 1$ and can be dropped from Eq. (7.110).

The open-circuit voltage gain in Eq. (7.109) can be expressed alternatively as

$$A_{vo} = -\frac{\alpha}{r_e} \frac{R_C}{1 + R_e/r_e}$$

Thus,

$$A_{vo} = -\frac{g_m R_C}{1 + R_e/r_e} \simeq -\frac{g_m R_C}{1 + g_m R_e} \tag{7.111}$$

Thus, including R_e reduces the voltage gain by the factor $(1 + g_m R_e)$, which is the same factor by which R_{in} is increased. This points out an interesting trade-off between gain and input resistance, a trade-off that the designer can exercise through the choice of an appropriate value for R_e.

The output resistance R_o can be found from the circuit in Fig. 7.38(b) by inspection:

$$R_o = R_C$$

If a load resistance R_L is connected at the amplifier output, A_v can be found as

$$A_v = A_{vo} \frac{R_L}{R_L + R_o}$$
$$= -\alpha \frac{R_C}{r_e + R_e} \frac{R_L}{R_L + R_C}$$
$$= -\alpha \frac{R_C \parallel R_L}{r_e + R_e} \tag{7.112}$$

which could have been written directly using Eq. (7.110). The overall voltage gain G_v can now be found:

$$G_v = \frac{R_{in}}{R_{in} + R_{sig}} \times -\alpha \frac{R_C \parallel R_L}{r_e + R_e}$$

Substituting for R_{in} from Eq. (7.107) and replacing α with $\beta/(\beta+1)$ results in

$$G_v = -\beta \frac{R_C \| R_L}{R_{sig} + (\beta+1)(r_e + R_e)} \tag{7.113}$$

Careful examination of this expression reveals that the denominator comprises the total resistance in the base circuit [recall that $(\beta+1)(r_e+R_e)$ is the reflection of (r_e+R_e) from the emitter side to the base side]. Thus the expression in Eq. (7.113) states that the voltage gain from base to collector is equal to β times the ratio of the total resistance in the collector to the total resistance in the base. The factor β appears because it is the ratio of the collector current to the base current. This general and useful expression has no counterpart in the MOS case. We observe that the overall voltage gain G_v is lower than the value without R_e, namely,

$$G_v = -\beta \frac{R_C \| R_L}{R_{sig} + (\beta+1)r_e} \tag{7.114}$$

because of the additional term $(\beta+1)R_e$ in the denominator. The gain, however, is now less sensitive to the value of β, a desirable result because of the typical wide variability in the value of β.

Another important consequence of including the resistance R_e in the emitter is that it enables the amplifier to handle larger input signals without incurring nonlinear distortion. This is because only a fraction of the input signal at the base, v_i, appears between the base and the emitter. Specifically, from the circuit in Fig. 7.38(b), we see that

$$\frac{v_\pi}{v_i} = \frac{r_e}{r_e + R_e} \simeq \frac{1}{1 + g_m R_e} \tag{7.115}$$

Thus, for the same v_π, the signal at the input terminal of the amplifier, v_i, can be greater than for the CE amplifier by the factor $(1+g_mR_e)$.

To summarize, including a resistance R_e in the emitter of the CE amplifier results in the following characteristics:

1. The input resistance R_{in} is increased by the factor $(1+g_mR_e)$.
2. The voltage gain from base to collector, A_v, is reduced by the factor $(1+g_mR_e)$.
3. For the same nonlinear distortion, the input signal v_i can be increased by the factor $(1+g_mR_e)$.
4. The overall voltage gain is less dependent on the value of β.
5. The high-frequency response is significantly improved (as we shall see in Chapter 10).

With the exception of gain reduction, these characteristics represent performance improvements. Indeed, the reduction in gain is the price paid for obtaining the other performance improvements. In many cases this is a good bargain; it is the underlying philosophy for the use of negative feedback. That the resistance R_e introduces negative feedback in the amplifier circuit can be verified by utilizing a procedure similar to that we used above for the MOSFET case. In Chapter 11, where we shall study negative feedback formally, we will find that the factor $(1+g_mR_e)$, which appears repeatedly, is the "amount of negative feedback" introduced by R_e. Finally, we note that the negative-feedback action of R_e gives it the name **emitter degeneration resistance**.

Example 7.9

For the CE amplifier specified in Example 7.8, what value of R_e is needed to raise R_{in} to a value four times that of R_{sig}? With R_e included, find A_{vo}, R_o, A_v, and G_v. Also, if $\hat{v}_\pi$ is limited to 5 mV, what are the corresponding values of $\hat{v}_{sig}$ and $\hat{v}_o$?

Solution

To obtain $R_{in} = 4R_{sig} = 4 \times 5 = 20$ kΩ, the required R_e is found from

$$20 = (\beta + 1)(r_e + R_e)$$

With $\beta = 100$,

$$r_e + R_e \simeq 200\ \Omega$$

Thus,

$$R_e = 200 - 25 = 175\ \Omega$$

$$A_{vo} = -\alpha \frac{R_C}{r_e + R_e}$$

$$\simeq -\frac{5000}{25 + 175} = -25\ \text{V/V}$$

$$R_o = R_C = 5\ \text{k}\Omega\ \text{(unchanged)}$$

$$A_v = A_{vo} \frac{R_L}{R_L + R_o} = -25 \times \frac{5}{5+5} = -12.5\ \text{V/V}$$

$$G_v = \frac{R_{in}}{R_{in} + R_{sig}} A_v = -\frac{20}{20+5} \times 12.5 = -10\ \text{V/V}$$

For $\hat{v}_\pi = 5$ mV,

$$\hat{v}_i = \hat{v}_\pi \left(\frac{r_e + R_e}{r_e} \right)$$

$$= 5\left(1 + \frac{175}{25}\right) = 40\ \text{mV}$$

$$\hat{v}_{sig} = \hat{v}_i \frac{R_{in} + R_{sig}}{R_{in}}$$

$$= 40\left(1 + \frac{5}{20}\right) = 50\ \text{mV}$$

$$\hat{v}_o = \hat{v}_{sig} \times |G_v|$$

$$= 50 \times 10 = 500\ \text{mV} = 0.5\ \text{V}$$

Thus, while $|G_v|$ has decreased to about a third of its original value, the amplifier is able to produce as large an output signal as before for the same nonlinear distortion.

EXERCISE

7.24 Show that with R_e included, and v_π limited to a maximum value $\hat{v}_\pi$, the maximum allowable input signal, $\hat{v}_{sig}$, is given by

$$\hat{v}_{sig} = \hat{v}_\pi \left(1 + \frac{R_e}{r_e} + \frac{R_{sig}}{r_\pi}\right)$$

If the transistor is biased at $I_C = 0.5$ mA and has a β of 100, what value of R_e is needed to permit an input signal $\hat{v}_{sig}$ of 100 mV from a source with a resistance $R_{sig} = 10$ kΩ while limiting $\hat{v}_\pi$ to 10 mV? What is R_{in} for this amplifier? If the total resistance in the collector is 10 kΩ, what G_v value results?
Ans. 350 Ω; 40.4 kΩ; −19.8 V/V

7.3.5 The Common-Gate (CG) and the Common-Base (CB) Amplifiers

Figure 7.39(a) shows a common-gate amplifier with the biasing circuit omitted. The amplifier is fed with a signal source characterized by v_{sig} and R_{sig}. Since R_{sig} appears in series with the source, it is more convenient to represent the transistor with the T model than with the π model. Doing this, we obtain the amplifier equivalent circuit shown in Fig. 7.39(b).

From inspection of the equivalent circuit of Fig. 7.39(b), we see that the input resistance

$$R_{in} = \frac{1}{g_m} \tag{7.116}$$

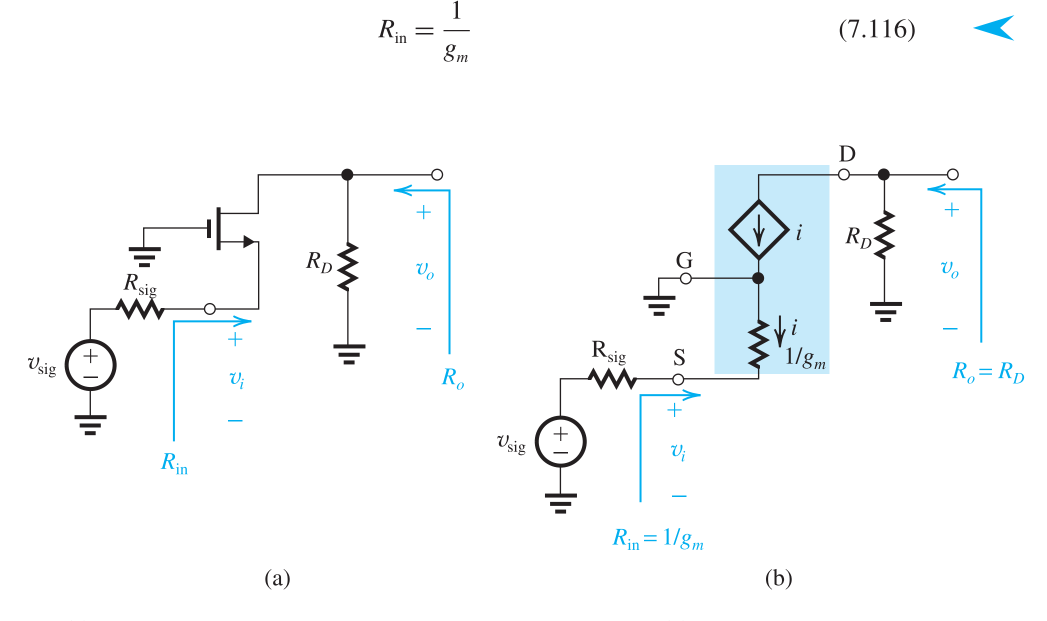


Figure 7.39 **(a)** Common-gate (CG) amplifier with bias arrangement omitted. **(b)** Equivalent circuit of the CG amplifier with the MOSFET replaced with its T model.

This should have been expected, since we are looking into the source and the gate is grounded. Typically $1/g_m$ is a few hundred ohms; thus the CG amplifier has a low input resistance.

To determine the voltage gain A_{vo}, we write at the drain node

$$v_o = -iR_D$$

and substitute for the source current i from

$$i = -\frac{v_i}{1/g_m}$$

to obtain

$$A_{vo} \equiv \frac{v_o}{v_i} = g_m R_D \tag{7.117}$$

which except for the positive sign is identical to the expression for A_{vo} of the CS amplifier.

The output resistance of the CG circuit can be found by inspection of the circuit in Fig. 7.39(b) as

$$R_o = R_D \tag{7.118}$$

which is the same as in the case of the CS amplifier.

Although the gain of the CG amplifier proper has the same magnitude as that of the CS amplifier, this is usually not the case as far as the overall voltage gain is concerned. The low input resistance of the CG amplifier can cause the input signal to be severely attenuated. Specifically,

$$\frac{v_i}{v_{sig}} = \frac{R_{in}}{R_{in} + R_{sig}} = \frac{1/g_m}{1/g_m + R_{sig}} \tag{7.119}$$

from which we see that except for situations in which R_{sig} is on the order of $1/g_m$, the signal transmission factor v_i/v_{sig} can be very small and the overall voltage gain G_v can be correspondingly small. Specifically, with a resistance R_L connected at the output

$$G_v = \frac{1/g_m}{R_{sig} + 1/g_m}[g_m(R_D \| R_L)]$$

Thus,

$$G_v = \frac{(R_D \| R_L)}{R_{sig} + 1/g_m} \tag{7.120}$$

Observe that *the overall voltage gain is simply the ratio of the total resistance in the drain circuit to the total resistance in the source circuit*. If R_{sig} is of the same order as R_D and R_L, G_v will be very small.

Because of its low input resistance, the CG amplifier alone has very limited application. One such application is to amplify high-frequency signals that come from sources with relatively low resistances. These include cables, where it is usually necessary for the input resistance of the amplifier to match the characteristic resistance of the cable. As will be shown in Chapter 10, the CG amplifier has excellent high-frequency response. Thus it can be combined with the CS amplifier in a very beneficial way that takes advantage of the best features of each of the two configurations. A very significant circuit of this kind will be studied in Chapter 8.

EXERCISE

7.25 A CG amplifier is required to match a signal source with $R_{sig} = 100\ \Omega$. At what current I_D should the MOSFET be biased if it is operated at an overdrive voltage of 0.20 V? If the total resistance in the drain circuit is 2 kΩ, what overall voltage gain is realized?
Ans. 1 mA; 10 V/V

Very similar results can be obtained for the CB amplifier shown in Fig. 7.40(a). Specifically, from the equivalent circuit in Fig. 7.40(b) we can find

$$R_{in} = r_e = \frac{\alpha}{g_m} \simeq 1/g_m \tag{7.121}$$

$$A_{vo} = \frac{\alpha}{r_e} R_C = g_m R_C \tag{7.122}$$

$$R_o = R_C \tag{7.123}$$

and with a load resistance R_L connected to the output, the overall voltage gain is given by

$$G_v \equiv \frac{v_o}{v_{sig}} = \alpha \frac{R_C \| R_L}{R_{sig} + r_e} \tag{7.124}$$

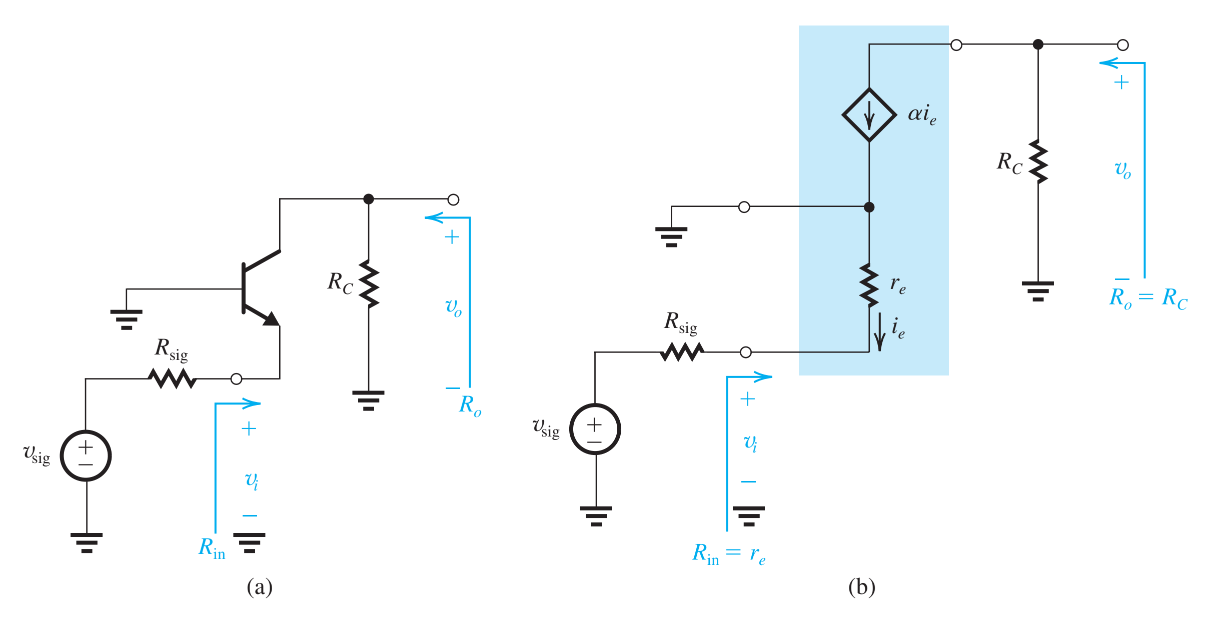


Figure 7.40 (**a**) CB amplifier with bias details omitted; (**b**) amplifier equivalent circuit with the BJT represented by its T model.

Since $\alpha \simeq 1$, we see that as in the case of the CG amplifier, the overall voltage gain is simply the ratio of the total resistance in the collector to the total resistance in the emitter. We also note that the overall voltage gain is almost independent of the value of β (except through the small dependence of α on β), a desirable property. Observe that for R_{sig} of the same order as R_C and R_L, the gain will be very small.

In summary, the CB and CG amplifiers exhibit a very low input resistance ($1/g_m$), an open-circuit voltage gain that is positive and equal in magnitude to that of the CE (CG) amplifier (g_mR_C or g_mR_D), and, like the CE (CS) amplifier, a relatively high output resistance (R_C or R_D). Because of its very low input resistance, the CB (CG) circuit *alone* is not attractive as a voltage amplifier except in specialized applications, such as the cable amplifier mentioned above. The CB (CG) amplifier has excellent high-frequency performance, which as we shall see in Chapters 8 and 10, makes it useful in combination with other circuits in the implementation of high-frequency amplifiers.

EXERCISES

7.26 Consider a CB amplifier utilizing a BJT biased at $I_C = 1$ mA and with $R_C = 5$ kΩ. Determine R_{in}, A_{vo}, and R_o. If the amplifier is loaded in $R_L = 5$ kΩ, what value of A_v results? What G_v is obtained if $R_{sig} = 5$ kΩ?
Ans. 25 Ω; 200 V/V; 5 kΩ; 100 V/V; 0.5 V/V

7.27 A CB amplifier is required to amplify a signal delivered by a coaxial cable having a characteristic resistance of 50 Ω. What bias current I_C should be utilized to obtain R_{in} that is matched to the cable resistance? To obtain an overall voltage gain of G_v of 40 V/V, what should the total resistance in the collector (i.e., $R_C \| R_L$) be?
Ans. 0.5 mA; 4 kΩ

7.3.6 The Source and Emitter Followers

The last of the basic transistor amplifier configurations is the common-drain (common-collector) amplifier, an important circuit that finds application in the design of both small-signal amplifiers and amplifiers that are required to handle large signals and deliver substantial amounts of signal power to a load. This latter variety will be studied in Chapter 12. The common-drain amplifier is more commonly known as the *source follower*, and the common-collector amplifier is more commonly known as the *emitter follower*. The reason behind these names will become apparent shortly.

The Need for Voltage Buffers Before embarking on the analysis of the source and the emitter followers, it is useful to look at one of their more common applications. Consider the situation depicted in Fig. 7.41(a). A signal source delivering a signal of reasonable strength (1 V) with an internal resistance of 1 MΩ is to be connected to a 1-kΩ load resistance. Connecting the source to the load directly as in Fig. 7.41(b) would result in severe attenuation

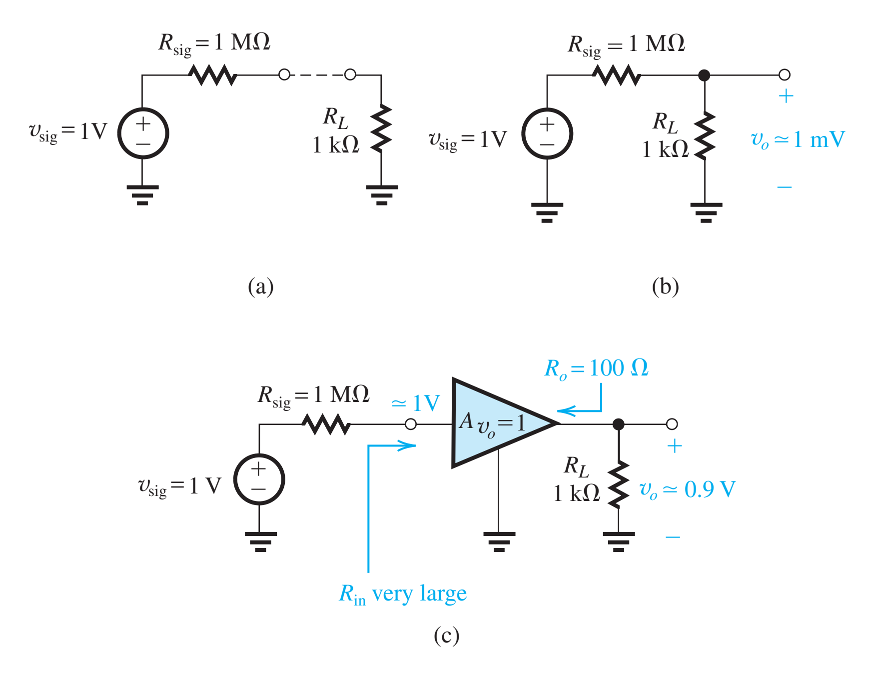


Figure 7.41 Illustrating the need for a unity-gain voltage buffer amplifier.

of the signal; the signal appearing across the load will be only 1/(1000 + 1) of the input signal, or about 1 mV. An alternative course of action is suggested in Fig. 7.41(c). Here we have interposed an amplifier between the source and the load. Our amplifier, however, is unlike the amplifiers we have been studying in this chapter thus far; it has a voltage gain of only unity. This is because our signal is already of sufficient strength and we do not need to increase its amplitude. Note, however, that our amplifier has a very high input resistance, thus almost all of v_{sig} (i.e., 1 V) will appear at the input of the amplifier proper. Since the amplifier has a low output resistance (100 Ω), 90% of this signal (0.9 V) will appear at the output, obviously a very significant improvement over the situation without the amplifier. As will be seen next, the source follower can easily implement the unity-gain buffer amplifier shown in Fig. 7.41(c).

Characteristic Parameters of the Source Follower Figure 7.42(a) shows a source follower with the bias circuit omitted. The source follower is fed with a signal generator (v_{sig}, R_{sig}) and has a load resistance R_L connected between the source terminal and ground. We shall assume that R_L includes both the actual load and any other resistance that may be present between the source terminal and ground (e.g., for biasing purposes). Normally, the actual load resistance would be much lower in value than such other resistances and thus would dominate.

Since the MOSFET has a resistance R_L connected in its source terminal, it is most convenient to use the T model, as shown in Fig. 7.40(b). From the latter circuit we can write by inspection

$$R_{in} = \infty$$

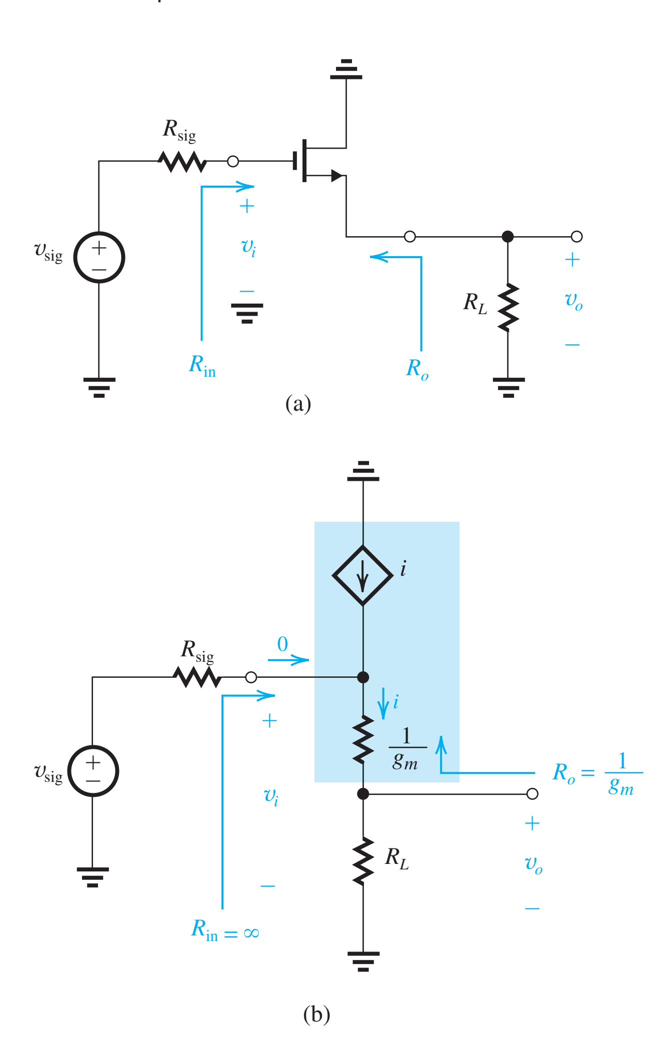


Figure 7.42 **(a)** Common-drain amplifier or source follower with the bias circuit omitted. **(b)** Equivalent circuit of the source follower obtained by replacing the MOSFET with its T model.

and obtain A_v from the voltage divider formed by $1/g_m$ and R_L as

$$A_v \equiv \frac{v_o}{v_i} = \frac{R_L}{R_L + 1/g_m} \tag{7.125}$$

Setting $R_L = \infty$ we obtain

$$A_{vo} = 1 \tag{7.126}$$

The output resistance R_o is found by setting $v_i = 0$ (i.e., by grounding the gate). Now looking back into the output terminal, excluding R_L, we simply see $1/g_m$, thus

$$R_o = 1/g_m \tag{7.127}$$

The unity open-circuit voltage gain together with R_o in Eq. (7.127) can be used to find A_v when a load resistance R_L is connected. The result is simply the expression in Eq. (7.125).

Finally, because of the infinite R_{in}, $v_i = v_{sig}$, and the overall voltage gain is

$$G_v = A_v = \frac{R_L}{R_L + 1/g_m} \tag{7.128}$$

Thus G_v will be lower than unity. However, because $1/g_m$ is usually low, the voltage gain can be close to unity. The unity open-circuit voltage gain in Eq. (7.126) indicates that the voltage at the source terminal will follow that at the input, hence the name *source follower*.

In conclusion, the source follower features a very high input resistance (ideally, infinite), a relatively low output resistance ($1/g_m$), and an open-circuit voltage gain that is near unity (ideally, unity). Thus the source follower is ideally suited for implementing the unity-gain voltage buffer of Fig. 7.41(c). The source follower is also used as the output (i.e., last) stage in a multistage amplifier, where its function is to equip the overall amplifier with a low output resistance, thus enabling it to supply relatively large load currents without loss of gain (i.e., with little reduction of output signal level). The design of output stages is studied in Chapter 12.

EXERCISES

D7.28 It is required to design a source follower that implements the buffer amplifier shown in Fig. 7.41(c). If the MOSFET is operated with an overdrive voltage $V_{OV} = 0.25$ V, at what drain current should it be biased? Find the output signal amplitude and the signal amplitude between gate and source.
Ans. 1.25 mA; 0.91 V; 91 mV

D7.29 A MOSFET is connected in the source-follower configuration and employed as the output stage of a cascade amplifier. It is required to provide an output resistance of 200 Ω. If the MOSFET has $k_n' = 0.4$ mA/V^2 and is operated at $V_{OV} = 0.25$ V, find the required W/L ratio. Also specify the dc bias current I_D. If the amplifier load resistance varies over the range 1 kΩ to 10 kΩ, what is the range of G_v of the source follower?
Ans. 50; 0.625 mA; 0.83 V/V to 0.98 V/V

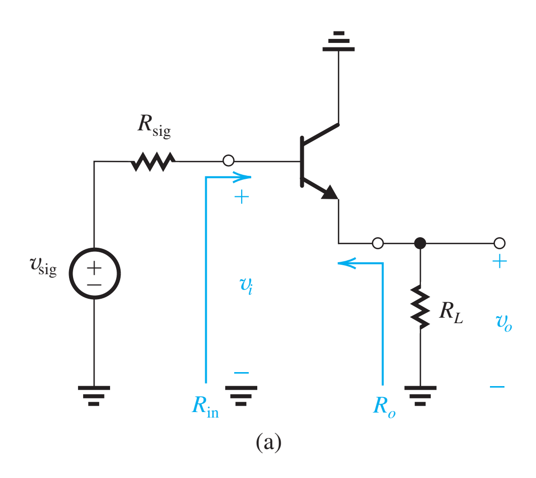
Characteristic Parameters of the Emitter Follower Although the emitter follower does not have an infinite input resistance (as in the case of the source follower), it is still widely used as a voltage buffer. In fact, it is a very versatile and popular circuit. We will therefore study it in some detail.

Figure 7.43(a) shows an emitter follower with the equivalent circuit shown in Fig. 7.43(b). The input resistance R_{in} is found from

$$R_{in} = \frac{v_i}{i_b}$$

Substituting for $i_b = i_e/(\beta + 1)$ where i_e is given by

$$i_e = \frac{v_i}{r_e + R_L}$$



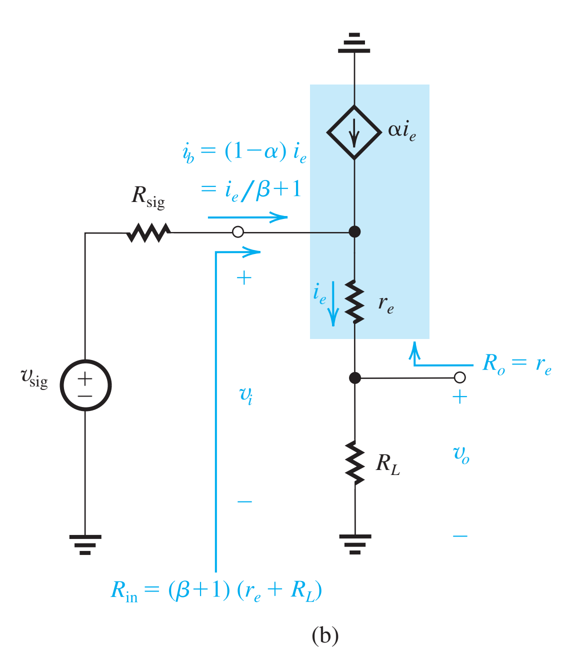


Figure 7.43 (**a**) Common-collector amplifier or emitter follower with the bias circuit omitted. (**b**) Equivalent circuit obtained by replacing the BJT with its T model.

we obtain

$$R_{in} = (\beta+1)(r_e + R_L) \tag{7.129}$$

a result that we could have written directly, utilizing the resistance-reflection rule. Note that as expected the emitter follower takes the low load resistance and reflects it to the base side, where the signal source is, after increasing its value by a factor $(\beta+1)$. It is this impedance transformation property of the emitter follower that makes it useful in

connecting a low-resistance load to a high-resistance source, that is, to implement a buffer amplifier.

The voltage gain A_v is given by

$$A_v \equiv \frac{v_o}{v_i} = \frac{R_L}{R_L + r_e} \tag{7.130}$$

Setting $R_L = \infty$ yields A_{vo},

$$A_{vo} = 1 \tag{7.131}$$

Thus, as expected, the open-circuit voltage gain of the emitter follower proper is unity, which means that the signal voltage at the emitter *follows* that at the base, which is the origin of the name "emitter follower."

To determine R_o, refer to Fig. 7.43(b) and look back into the emitter (i.e., behind or excluding R_L) while setting $v_i = 0$ (i.e., grounding the base). You will see r_e of the BJT, thus

$$R_o = r_e \tag{7.132}$$

This result together with $A_{vo} = 1$ yields A_v in Eq. (7.130), thus confirming our earlier analysis.

We next determine the overall voltage gain G_v, as follows:

$$\frac{v_i}{v_{sig}} = \frac{R_{in}}{R_{in} + R_{sig}}$$

$$= \frac{(\beta + 1)(r_e + R_L)}{(\beta + 1)(r_e + R_L) + R_{sig}}$$

$$G_v \equiv \frac{v_o}{v_{sig}} = \frac{v_i}{v_{sig}} \times A_v$$

Substituting for A_v from Eq. (7.130) results in

$$G_v = \frac{(\beta + 1)R_L}{(\beta + 1)R_L + (\beta + 1)r_e + R_{sig}} \tag{7.133}$$

This equation indicates that the overall gain, though lower than one, can be close to one if $(\beta + 1)R_L$ is larger or comparable in value to R_{sig}. This again confirms the action of the emitter follower in delivering a large proportion of v_{sig} to a low-valued load resistance R_L even though R_{sig} can be much larger than R_L. The key point is that R_L is multiplied by $(\beta + 1)$ before it is "presented to the source." Figure 7.44(a) shows an equivalent circuit of the emitter follower obtained by simply reflecting r_e and R_L to the base side. The overall voltage gain $G_v \equiv v_o/v_{sig}$ can be determined directly and very simply from this circuit by using the voltage divider rule. The result is the expression for G_v already given in Eq. (7.133).

Dividing all resistances in the circuit of Fig. 7.44(a) by $\beta + 1$ does not change the voltage ratio v_o/v_{sig}. Thus we obtain another equivalent circuit, shown in Fig. 7.44(b), that can be used to determine $G_v \equiv v_o/v_{sig}$ of the emitter follower. A glance at this circuit reveals that it is simply the equivalent circuit obtained by reflecting v_{sig} and R_{sig} from the base side to the emitter side. In this reflection, v_{sig} does not change, but R_{sig} is divided by $\beta + 1$. Thus, we

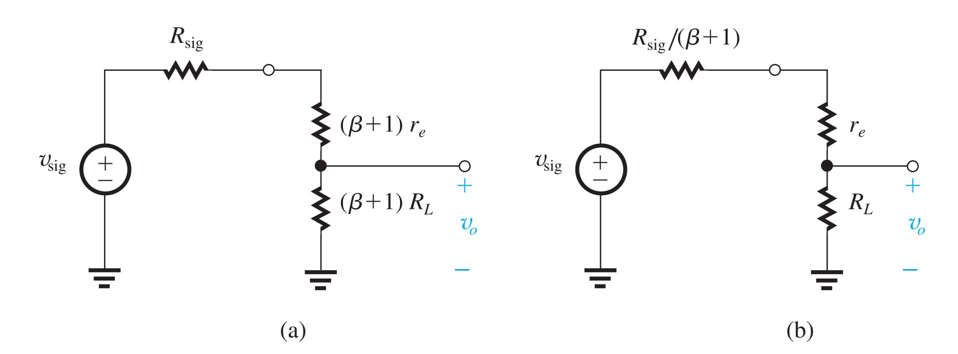


Figure 7.44 Simple equivalent circuits for the emitter follower obtained by (**a**) reflecting r_e and R_L to the base side, and (**b**) reflecting v_{sig} and R_{sig} to the emitter side. Note that the circuit in (**b**) can be obtained from that in (**a**) by simply dividing all resistances by $(\beta + 1)$.

either reflect to the base side and obtain the circuit in Fig. 7.44(a) or reflect to the emitter side and obtain the circuit in Fig. 7.44(b). From the latter, G_v can be found as

$$G_v \equiv \frac{v_o}{v_{sig}} = \frac{R_L}{R_L + r_e + R_{sig}/(\beta+1)} \tag{7.134}$$

Observe that this expression is the same as that in Eq. (7.133) except for dividing both the numerator and denominator by $\beta + 1$.

The expression for G_v in Eq. (7.134) has an interesting interpretation: The emitter follower reduces R_{sig} by the factor $(\beta + 1)$ before "presenting it to the load resistance R_L" : an impedance transformation that has the same buffering effect.

At this point it is important to note that although the emitter follower does not provide voltage gain it has a current gain of $\beta + 1$.

Thévenin Representation of the Emitter-Follower Output A more general representation of the emitter-follower output is shown in Fig. 7.45(a). Here G_{vo} is the overall open-circuit voltage gain that can be obtained by setting $R_L = \infty$ in the circuit of Fig. 7.44(b), as illustrated in Fig. 7.45(b). The result is $G_{vo} = 1$. The output resistance R_{out} is *different* from R_o. To determine R_{out} we set v_{sig} to zero (rather than setting v_i to zero). Again we can use the equivalent circuit in Fig. 7.44(b) to do this, as illustrated in Fig. 7.45(c). We see that

$$R_{out} = r_e + \frac{R_{sig}}{\beta+1} \tag{7.135}$$

Finally, we show in Fig. 7.45(d) the emitter-follower circuit together with its R_{in} and R_{out}. Observe that R_{in} is determined by reflecting r_e and R_L to the base side (by multiplying their values by $\beta + 1$). To determine R_{out}, grab hold of the emitter and walk (or just look!) backward while $v_{sig} = 0$. You will see r_e in series with R_{sig}, which because it is in the base must be divided by $(\beta + 1)$.

We note that unlike the amplifier circuits we studied earlier, the emitter follower is *not* unilateral. This is manifested by the fact that R_{in} depends on R_L and R_{out} depends on R_{sig}.

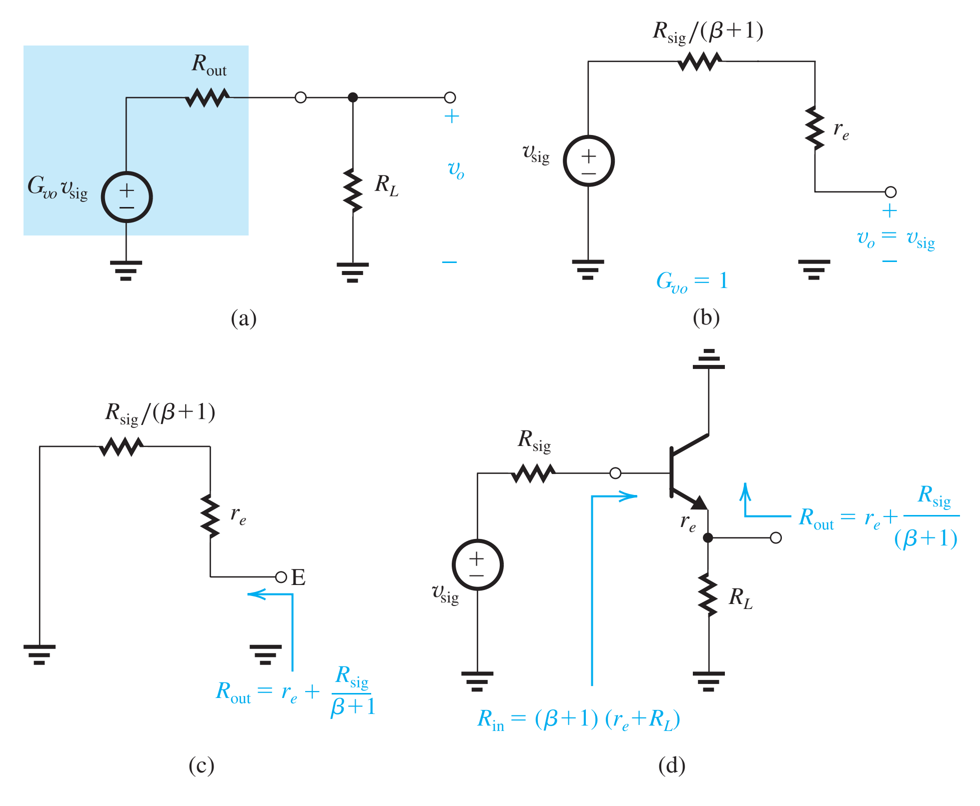


Figure 7.45 (**a**) Thévenin representation of the output of the emitter follower. (**b**) Obtaining G_{vo} from the equivalent circuit in Fig. 7.44(b). (**c**) Obtaining R_{out} from the equivalent circuit in Fig. 7.44(b) with v_{sig} set to zero. (**d**) The emitter follower with R_{in} and R_{out} determined simply by looking into the input and output terminals, respectively.

Example 7.10

It is required to design an emitter follower to implement the buffer amplifier of Fig. 7.46(a). Specify the required bias current I_E and the minimum value the transistor β must have. Determine the maximum allowed value of v_{sig} if v_π is to be limited to 5 mV in order to obtain reasonably linear operation. With $v_{sig} = 200$ mV, determine the signal voltage at the output if R_L is changed to 2 kΩ, and to 0.5 kΩ.

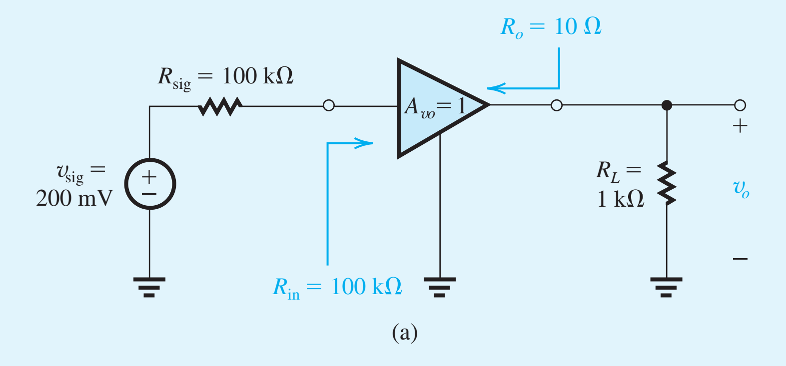


Figure 7.46 Circuit for Example 7.10.

Example 7.10 *continued*

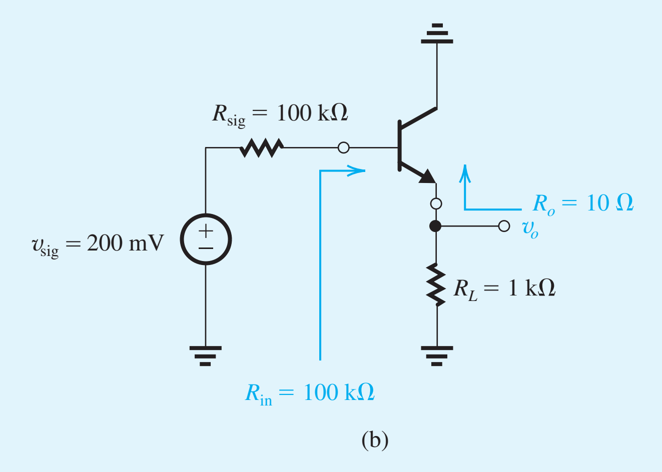


Figure 7.46 *continued*

Solution

The emitter-follower circuit is shown in Fig. 7.46(b). To obtain $R_o = 10\ \Omega$, we bias the transistor to obtain $r_e = 10\ \Omega$. Thus,

$$10\ \Omega = \frac{V_T}{I_E}$$
$$I_E = 2.5\text{ mA}$$

The input resistance R_{in} will be

$$R_{in} = (\beta + 1)(r_e + R_L)$$
$$100 = (\beta + 1)(0.01 + 1)$$

Thus, the BJT should have a β with a minimum value of 98. A higher β would obviously be beneficial.

The overall voltage gain can be determined from

$$G_v \equiv \frac{v_o}{v_{sig}} = \frac{R_L}{R_L + r_e + \dfrac{R_{sig}}{(\beta + 1)}}$$

Assuming $\beta = 100$, the value of G_v obtained is

$$G_v = 0.5$$

Thus when $v_{sig} = 200$ mV, the signal at the output will be 100 mV. Since the 100 mV appears across the 1-kΩ load, the signal across the base–emitter junction can be found from

$$v_\pi = \frac{v_o}{R_L} \times r_e$$
$$= \frac{100}{1000} \times 10 = 1 \text{ mV}$$

If $\hat{v}_\pi = 5$ mV then v_{sig} can be increased by a factor of 5, resulting in $\hat{v}_{sig} = 1$ V.

To obtain v_o as the load is varied, we use the Thévenin equivalent of the emitter follower, shown in Fig. 7.45(a) with $G_{vo} = 1$ and

$$R_{out} = \frac{R_{sig}}{\beta + 1} + r_e = \frac{100}{101} + 0.01 = 1 \text{ k}\Omega$$

to obtain

$$v_o = v_{sig} \frac{R_L}{R_L + R_{out}}$$

For $R_L = 2$ kΩ,

$$v_o = 200 \text{ mV} \times \frac{2}{2+1} = 133.3 \text{ mV}$$

and for $R_L = 0.5$ kΩ,

$$v_o = 200 \text{ mV} \times \frac{0.5}{0.5+1} = 66.7 \text{ mV}$$

EXERCISE

7.30 An emitter follower utilizes a transistor with $\beta = 100$ and is biased at $I_C = 5$ mA. It operates between a source having a resistance of 10 kΩ and a load of 1 kΩ. Find R_{in}, G_{vo}, R_{out}, and G_v. What is the peak amplitude of v_{sig} that results in v_π having a peak amplitude of 5 mV? Find the resulting peak amplitude at the output.

Ans. 101.5 kΩ; 1 V/V; 104 Ω; 0.91 V/V; 1.1 V; 1 V

7.3.7 Summary Tables and Comparisons

For easy reference and to enable comparisons, we present in Tables 7.4 and 7.5 the formulas for determining the characteristic parameters for the various configurations of MOSFET and BJT amplifiers, respectively. In addition to the remarks made throughout this section about the characteristics and areas of applicability of the various configurations, we make the following concluding points:

1. MOS amplifiers provide much higher, ideally infinite input resistances (except, of course, for the CG configuration). This is a definite advantage over BJT amplifiers.
2. BJTs exhibit higher g_m values than MOSFETs, resulting in higher gains.
3. For discrete-circuit amplifiers—that is, those that are assembled from discrete components on a printed-circuit board (PCB)—the BJT remains the device of choice. This is because discrete BJTs are much easier to handle physically than discrete MOSFETs and, more important, a very wide variety of discrete BJTs is available commercially. The remainder of this chapter is concerned with discrete-circuit amplifiers.
4. Integrated-circuit (IC) amplifiers predominantly use MOSFETs, although BJTs are utilized in certain niche areas. Chapters 8 to 13 are mainly concerned with IC amplifiers.
5. The CS and CE configurations are the best suited for realizing the bulk of the gain required in an amplifier. Depending on the magnitude of the gain required, either a single stage or a cascade of two or three stages can be used.
6. Including a resistance R_s in the source of the CS amplifier (a resistance R_e in the emitter of the CE amplifier) provides a number of performance improvements at the expense of gain reduction.

Table 7.4 Characteristics of MOSFET Amplifiers

	Characteristics[a]				
Amplifier type	R_{in}	A_{vo}	R_o	A_v	G_v
Common source (Fig. 7.35)	∞	$-g_m R_D$	R_D	$-g_m(R_D \parallel R_L)$	$-g_m(R_D \parallel R_L)$
Common source with R_s (Fig. 7.37)	∞	$-\dfrac{g_m R_D}{1+g_m R_s}$	R_D	$\dfrac{-g_m(R_D \parallel R_L)}{1+g_m R_s}$ $-\dfrac{R_D \parallel R_L}{1/g_m + R_s}$	$-\dfrac{g_m(R_D \parallel R_L)}{1+g_m R_s}$ $-\dfrac{R_D \parallel R_L}{1/g_m + R_s}$
Common gate (Fig. 7.39)	$\dfrac{1}{g_m}$	$g_m R_D$	R_D	$g_m(R_D \parallel R_L)$	$\dfrac{R_D \parallel R_L}{R_{sig} + 1/g_m}$
Source follower (Fig. 7.42)	∞	1	$\dfrac{1}{g_m}$	$\dfrac{R_L}{R_L + 1/g_m}$	$\dfrac{R_L}{R_L + 1/g_m}$

[a] For the interpretation of R_{in}, A_{vo}, and R_o, refer to Fig. 7.34(b).

Table 7.5 Characteristics of BJT Amplifiers[a,b]

	R_{in}	A_{vo}	R_o	A_v	G_v
Common emitter (Fig. 7.36)	$(\beta+1)r_e$	$-g_m R_C$	R_C	$-g_m(R_C \parallel R_L)$ $-\alpha \dfrac{R_C \parallel R_L}{r_e}$	$-\beta \dfrac{R_C \parallel R_L}{R_{sig}+(\beta+1)r_e}$
Common emitter with R_e (Fig. 7.38)	$(\beta+1)(r_e+R_e)$	$-\dfrac{g_m R_C}{1+g_m R_e}$	R_C	$\dfrac{-g_m(R_C \parallel R_L)}{1+g_m R_e}$ $-\alpha \dfrac{R_C \parallel R_L}{r_e+R_e}$	$-\beta \dfrac{R_C \parallel R_L}{R_{sig}+(\beta+1)(r_e+R_e)}$
Common base (Fig. 7.40)	r_e	$g_m R_C$	R_C	$g_m(R_C \parallel R_L)$ $\alpha \dfrac{R_C \parallel R_L}{r_e}$	$\alpha \dfrac{R_C \parallel R_L}{R_{sig}+r_e}$
Emitter follower (Fig. 7.43)	$(\beta+1)(r_e+R_L)$	1	r_e	$\dfrac{R_L}{R_L+r_e}$	$\dfrac{R_L}{R_L+r_e+R_{sig}/(\beta+1)}$ $G_{vo}=1$ $R_{out}=r_e+\dfrac{R_{sig}}{\beta+1}$

[a] For the interpretation of R_m, A_{vo}, and R_o refer to Fig. 7.34.
[b] Setting $\beta=\infty$ $(\alpha=1)$ and replacing r_e with $1/g_m$, R_C with R_D, and R_e with R_s results in the corresponding formulas for MOSFET amplifiers (Table 7.4).

7. The low input resistance of the CG and CB amplifiers makes them useful only in specific applications. As we shall see in Chapter 10, these two configurations exhibit a much better high-frequency response than that available from the CS and CE amplifiers. This makes them useful as high-frequency amplifiers, especially when combined with the CS or CE circuit. We shall study one such combination in Chapter 8.
8. The source follower (emitter follower) finds application as a voltage buffer for connecting a high-resistance source to a low-resistance load, and as the output stage in a multistage amplifier, where its purpose is to equip the amplifier with a low output resistance.

7.3.8 When and How to Include the Output Resistance r_o

So far we have been neglecting the output resistance r_o of the MOSFET and the BJT. We have done this for two reasons:

1. To keep things simple and focus attention on the significant features of each of the basic configurations, and
2. Because our main interest in this chapter is discrete-circuit design, where the circuit resistances (e.g., R_C, R_D, and R_L) are usually much smaller than r_o.

LEE DE FOREST—A FATHER OF THE ELECTRONICS AGE:

In 1906 self-employed inventor Lee de Forest (1873–1961) created a three-terminal vacuum tube; it was the first electronic amplifier of weak signals. The device was known initially as the de Forest valve. The patent filed in 1907, however, used the name Audion, with the "-ion" indicating that the device was not completely evacuated. By 1919, engineers had realized that complete evacuation of internal gases produced a more reliable device.

De Forest's first amplifier became known as the vacuum tube triode. Through its impact on radio, telephony, motion picture sound, and television, this invention, one of de Forest's 180 patents, is credited with introducing the electronics age. The vacuum tube, in a variety of types, remained the device for implementing amplifiers until the appearance of transistors in the early 1950s.

Nevertheless, in some instances it is relatively easy to include r_o in the analysis. Specifically:

1. In the CS and CE amplifiers, it can be seen that r_o of the transistor appears in parallel with R_D and R_C, respectively, and can be simply included in the corresponding formulas in Tables 7.4 and 7.5 by replacing R_D with $(R_D \| r_o)$ and R_C with $(R_C \| r_o)$. The effect will be a reduction in the magnitude of gain, of perhaps 5% to 10%.
2. In the source and emitter followers, it can be seen that the transistor r_o appears in parallel with R_L and can be taken into account by replacing R_L in the corresponding formulas with $(R_L \| r_o)$. Thus, here too, the effect of taking r_o into account is a small reduction in gain. More significant, however, taking r_o into account reduces the open-circuit voltage gain A_{vo} from unity to

$$A_{vo} = \frac{r_o}{r_o + (1/g_m)} \tag{7.136}$$

There are configurations in which taking r_o into account complicates the analysis considerably. These are the CS (CE) amplifiers with a source (emitter) resistance, and the CG (CB) amplifier. Fortunately, for discrete implementation of these configurations, the effect of neglecting r_o is usually small (which can be verified by computer simulation).

Finally, a very important point: *In the analysis and design of IC amplifiers, r_o must always be taken into account.* This is because, as will be seen in the next chapter, all the circuit resistances are of the same order of magnitude as r_o; thus, neglecting r_o can result in completely erroneous results.

7.4 Biasing

As discussed in Section 7.1, an essential step in the design of a transistor amplifier is the establishment of an appropriate dc operating point for the transistor. This is the step known as biasing or bias design. In this section, we study the biasing methods commonly employed in discrete-circuit amplifiers. Biasing of integrated-circuit amplifiers will be studied in Chapter 8.

Bias design aims to establish in the drain (collector) a dc current that is predictable and insensitive to variations in temperature and to the large variations in parameter values between devices of the same type. For instance, discrete BJTs belonging to the same manufacturer's part number can exhibit β values that range, say, from 50 to 150. Nevertheless, the bias design

for an amplifier utilizing this particular transistor type may specify that the dc collector current shall always be within, say, $\pm 10\%$ of the nominal value of, say, 1 mA. A similar statement can be made about the desired insensitivity of the dc drain current to the wide variations encountered in V_t of discrete MOSFETs.

A second consideration in bias design is locating the dc operating point in the active region of operation of the transistor so as to obtain high voltage gain while allowing for the required output signal swing without the transistor leaving the active region at any time (in order to avoid nonlinear distortion). We discussed this point in Section 7.1.7.

Although we shall consider the biasing of MOSFET and BJT amplifiers separately, the resulting circuits are very similar. Also, it will be seen that good bias designs incorporate a feedback mechanism that works to keep the dc bias point as constant as possible.

In order to keep matters simple and thus focus our attention on significant issues, we will neglect the Early effect; that is assume $\lambda = 0$ or $V_A = \infty$. This is certainly allowed in initial designs of discrete circuits. Of course, the design can be fine-tuned at a later point with the assistance of a circuit-simulation program such as SPICE.

7.4.1 The MOSFET Case

Biasing by Fixing V_{GS} The most straightforward approach to biasing a MOSFET is to fix its gate-to-source voltage V_{GS} to the value required[6] to provide the desired I_D. This voltage value can be derived from the power-supply voltage V_{DD} through the use of an appropriate voltage divider, as shown in Fig. 7.47(a). Alternatively, it can be derived from another suitable reference voltage that might be available in the system. Independent of how the voltage V_{GS} may be generated, this is *not* a good approach to biasing a MOSFET. To understand the reason for this statement, recall that

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_t)^2$$

and note that the values of the threshold voltage V_t, the oxide-capacitance C_{ox}, and (to a lesser extent) the transistor aspect ratio W/L vary widely among devices of supposedly the same size and type. This is certainly the case for discrete devices, in which large spreads in the values of these parameters occur among devices of the same manufacturer's part number. The spread can also be large in integrated circuits, especially among devices fabricated on different wafers and certainly between different batches of wafers. Furthermore, both V_t and μ_n depend on temperature, with the result that if we fix the value of V_{GS}, the drain current I_D becomes very much temperature dependent.

To emphasize the point that biasing by fixing V_{GS} is not a good technique, we show in Fig. 7.47 two i_D–v_{GS} characteristic curves representing extreme values in a batch of MOSFETs of the same type. Observe that for the fixed value of V_{GS}, the resultant spread in the values of the drain current can be substantial.

Biasing by Fixing V_G and Connecting a Resistance in the Source An excellent biasing technique for discrete MOSFET circuits consists of fixing the dc voltage at the gate, V_G, and connecting a resistance in the source lead, as shown in Fig. 7.48(a). For this circuit

[6]That is indeed what we were doing in Section 7.1. However, the amplifier circuits studied there were conceptual ones, not actual practical circuits. Our purpose in this section is to study the latter.

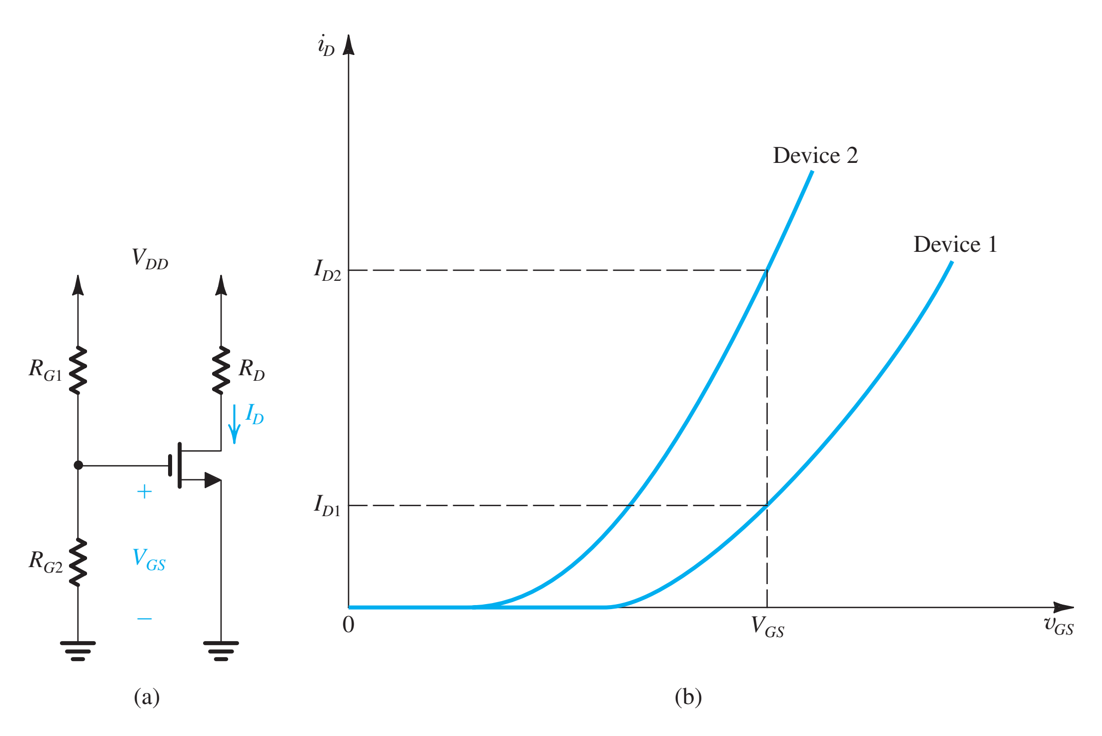


Figure 7.47 (**a**) Biasing the MOSFET with a constant V_{GS} generated from V_{DD} using a voltage divider (R_{G1}, R_{G2}); (**b**) the use of fixed bias (constant V_{GS}) can result in a large variability in the value of I_D. Devices 1 and 2 represent extremes among units of the same type.

we can write

$$V_G = V_{GS} + R_S I_D \tag{7.137}$$

Now, if V_G is much greater than V_{GS}, I_D will be mostly determined by the values of V_G and R_S. However, even if V_G is not much larger than V_{GS}, resistor R_S provides *negative feedback,* which acts to stabilize the value of the bias current I_D. To see how this comes about, consider what happens when I_D increases for whatever reason. Equation (7.137) indicates that since V_G is constant, V_{GS} will have to decrease. This in turn results in a decrease in I_D, a change that is opposite to that initially assumed. Thus the action of R_S works to keep I_D as constant as possible.[7]

Figure 7.48(b) provides a graphical illustration of the effectiveness of this biasing scheme. Here too we show the i_D–v_{GS} characteristics for two devices that represent the extremes of a batch of MOSFETs. Superimposed on the device characteristics is a straight line that represents the constraint imposed by the bias circuit—namely, Eq. (7.137). The intersection of this straight line with the i_D–v_{GS} characteristic curve provides the coordinates (I_D and V_{GS}) of the bias point. Observe that compared to the case of fixed V_{GS}, here the variability obtained in I_D is much smaller. Also, note that the variability decreases as V_G and R_S are made larger (thus providing a bias line that is less steep).

Two possible practical discrete implementations of this bias scheme are shown in Fig. 7.48(c) and (e). The circuit in Fig. 7.48(c) utilizes one power-supply V_{DD} and derives V_G

[7]The action of R_S in stabilizing the value of the bias current I_D is not unlike that of the resistance R_s, which we included in the source lead of a CS amplifier in Section 7.3.4. In the latter case also, R_s works to reduce the change in i_D with the result that the amplifier gain is reduced.

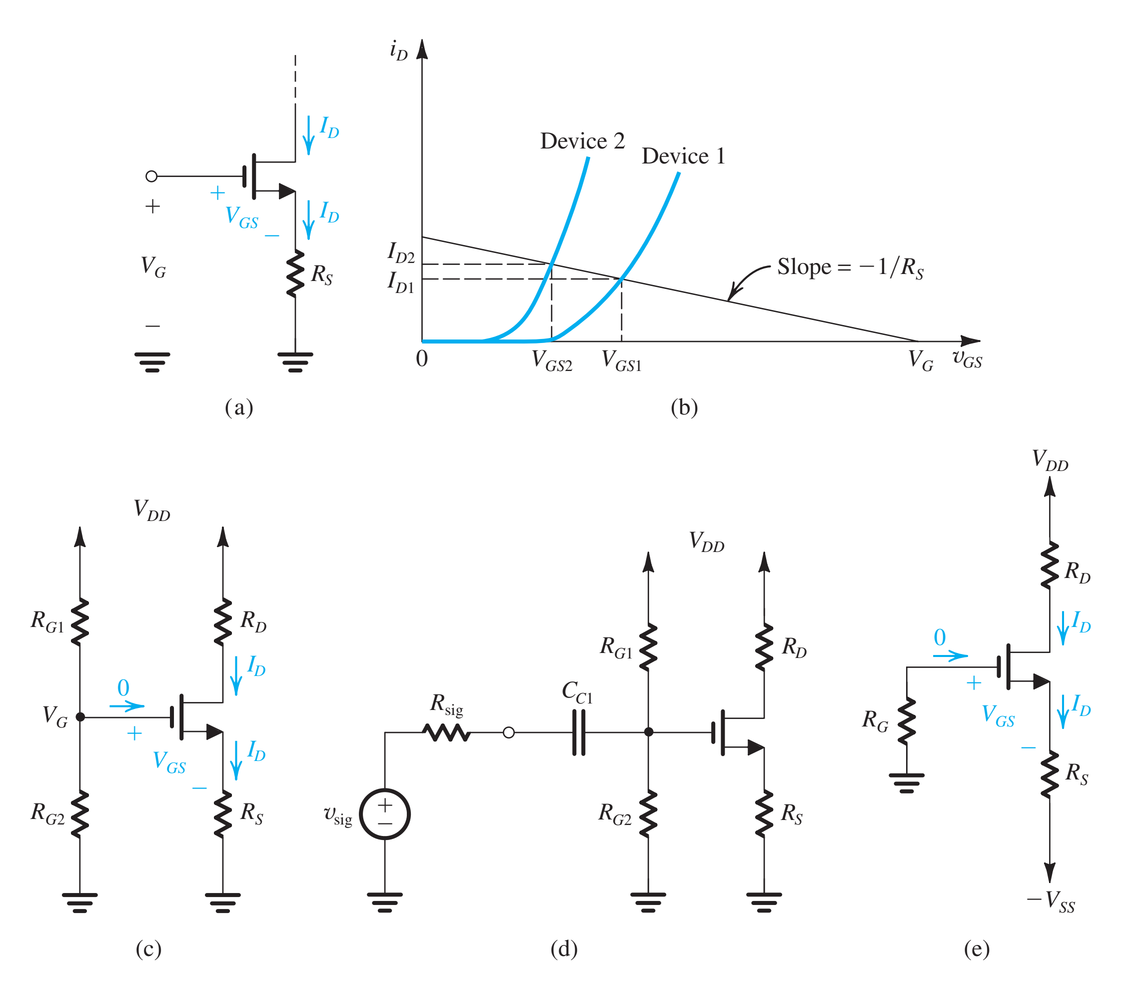


Figure 7.48 Biasing using a fixed voltage at the gate, V_G, and a resistance in the source lead, R_S: **(a)** basic arrangement; **(b)** reduced variability in I_D; **(c)** practical implementation using a single supply; **(d)** coupling of a signal source to the gate using a capacitor C_{C1}; **(e)** practical implementation using two supplies.

through a voltage divider (R_{G1}, R_{G2}). Since $I_G = 0$, R_{G1} and R_{G2} can be selected to be very large (in the megohm range), allowing the MOSFET to present a large input resistance to a signal source that may be connected to the gate through a coupling capacitor, as shown in Fig. 7.48(d). Here capacitor C_{C1} blocks dc and thus allows us to couple the signal v_{sig} to the amplifier input without disturbing the MOSFET dc bias point. The value of C_{C1} should be selected large enough to approximate a short circuit at all signal frequencies of interest. We shall study capacitively coupled MOSFET amplifiers, which are suitable only in discrete-circuit design, in Section 7.5. Finally, note that in the circuit of Fig. 7.48(c), resistor R_D is selected to be as large as possible to obtain high gain but small enough to allow for the desired signal swing at the drain while keeping the MOSFET in saturation at all times.

When two power supplies are available, as is often the case, the somewhat simpler bias arrangement of Fig. 7.48(e) can be utilized. This circuit is an implementation of Eq. (7.137), with V_G replaced by V_{SS}. Resistor R_G establishes a dc ground at the gate and presents a high input resistance to a signal source that may be connected to the gate through a coupling capacitor.

Example 7.11

It is required to design the circuit of Fig. 7.48(c) to establish a dc drain current $I_D = 0.5$ mA. The MOSFET is specified to have $V_t = 1$ V and $k_n' W/L = 1$ mA/V^2. For simplicity, neglect the channel-length modulation effect (i.e., assume $\lambda = 0$). Use a power-supply $V_{DD} = 15$ V. Calculate the percentage change in the value of I_D obtained when the MOSFET is replaced with another unit having the same $k_n' W/L$ but $V_t = 1.5$ V.

Solution

As a rule of thumb for designing this classical biasing circuit, we choose R_D and R_S to provide one-third of the power-supply voltage V_{DD} as a drop across each of R_D, the transistor (i.e., V_{DS}), and R_S. For $V_{DD} = 15$ V, this choice makes $V_D = +10$ V and $V_S = +5$ V. Now, since I_D is required to be 0.5 mA, we can find the values of R_D and R_S as follows:

$$R_D = \frac{V_{DD} - V_D}{I_D} = \frac{15 - 10}{0.5} = 10\text{ k}\Omega$$

$$R_S = \frac{V_S}{R_S} = \frac{5}{0.5} = 10\text{ k}\Omega$$

The required value of V_{GS} can be determined by first calculating the overdrive voltage V_{OV} from

$$I_D = \frac{1}{2}k_n'(W/L)V_{OV}^2$$

$$0.5 = \frac{1}{2} \times 1 \times V_{OV}^2$$

which yields $V_{OV} = 1$ V, and thus,

$$V_{GS} = V_t + V_{OV} = 1 + 1 = 2\text{ V}$$

Now, since $V_S = +5$ V, V_G must be

$$V_G = V_S + V_{GS} = 5 + 2 = 7\text{ V}$$

To establish this voltage at the gate we may select $R_{G1} = 8$ MΩ and $R_{G2} = 7$ MΩ. The final circuit is shown in Fig. 7.49. Observe that the dc voltage at the drain (+10 V) allows for a positive signal swing of +5 V (i.e., up to V_{DD}) and a negative signal swing of 4 V [i.e., down to $(V_G - V_t)$].

If the NMOS transistor is replaced with another having $V_t = 1.5$ V, the new value of I_D can be found as follows:

$$I_D = \frac{1}{2} \times 1 \times \left(V_{GS} - 1.5\right)^2 \quad (7.138)$$

$$V_G = V_{GS} + I_D R_S$$

$$7 = V_{GS} + 10I_D \quad (7.139)$$

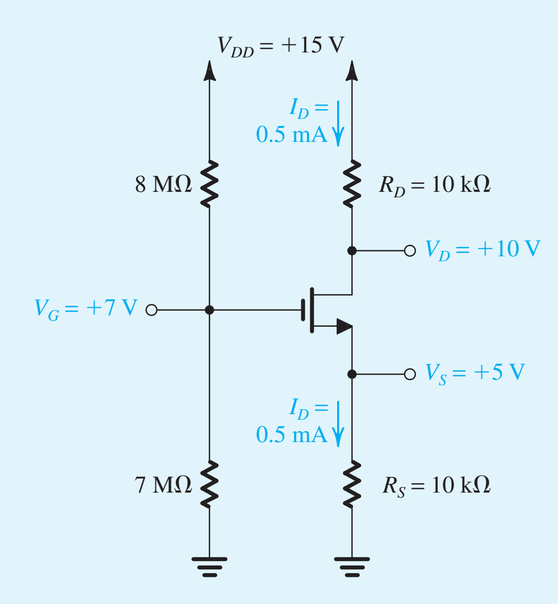


Figure 7.49 Circuit for Example 7.11.

Solving Eqs. (7.138) and (7.139) together yields

$$I_D = 0.455 \text{ mA}$$

Thus the change in I_D is

$$\Delta I_D = 0.455 - 0.5 = -0.045 \text{ mA}$$

which is $\dfrac{-0.045}{0.5} \times 100 = -9\%$ change.

EXERCISES

7.31 Consider the MOSFET in Example 7.11 when fixed-V_{GS} bias is used. Find the required value of V_{GS} to establish a dc bias current $I_D = 0.5$ mA. Recall that the device parameters are $V_t = 1$ V, $k_n' W/L = 1$ mA/V^2, and $\lambda = 0$. What is the percentage change in I_D obtained when the transistor is replaced with another having $V_t = 1.5$ V?
Ans. $V_{GS} = 2$ V; –75%

D7.32 Design the circuit of Fig. 7.48(e) to operate at a dc drain current of 0.5 mA and $V_D = +2$ V. Let $V_t = 1$ V, $k_n' W/L = 1$ mA/V^2, $\lambda = 0$, $V_{DD} = V_{SS} = 5$ V. Use standard 5% resistor values (see Appendix J), and give the resulting values of I_D, V_D, and V_S.
Ans. $R_D = R_S = 6.2$ kΩ; $I_D = 0.49$ mA, $V_S = -1.96$ V, and $V_D = +1.96$ V. R_G can be selected in the range of 1 MΩ to 10 MΩ.

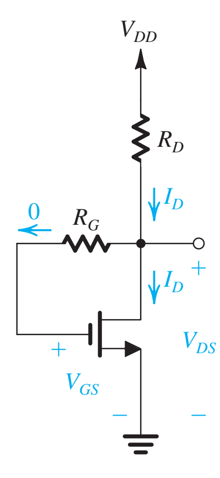


Figure 7.50 Biasing the MOSFET using a large drain-to-gate feedback resistance, R_G.

Biasing Using a Drain-to-Gate Feedback Resistor A simple and effective discrete-circuit biasing arrangement utilizing a feedback resistor connected between the drain and the gate is shown in Fig. 7.50. Here the large feedback resistance R_G (usually in the megohm range) forces the dc voltage at the gate to be equal to that at the drain (because $I_G = 0$). Thus we can write

$$V_{GS} = V_{DS} = V_{DD} - R_D I_D$$

which can be rewritten in the form

$$V_{DD} = V_{GS} + R_D I_D \tag{7.140}$$

which is identical in form to Eq. (7.137), which describes the operation of the bias scheme discussed above [that in Fig. 7.48(a)]. Thus, here too, if I_D for some reason changes, say increases, then Eq. (7.140) indicates that V_{GS} must decrease. The decrease in V_{GS} in turn causes a decrease in I_D, a change that is opposite in direction to the one originally assumed. Thus the negative feedback or degeneration provided by R_G works to keep the value of I_D as constant as possible.

The circuit of Fig. 7.50 can be utilized as an amplifier by applying the input voltage signal to the gate via a coupling capacitor so as not to disturb the dc bias conditions already established. The amplified output signal at the drain can be coupled to another part of the circuit, again via a capacitor. We considered such an amplifier circuit in Section 7.2 (Example 7.3).

EXERCISE

D7.33 Design the circuit in Fig. 7.50 to operate at a dc drain current of 0.5 mA. Assume $V_{DD} = +5$ V, $k_n' W/L = 1$ mA/V^2, $V_t = 1$ V, and $\lambda = 0$. Use a standard 5% resistance value for R_D, and give the actual values obtained for I_D and V_D.
Ans. $R_D = 6.2$ kΩ; $I_D \simeq 0.49$ mA; $V_D \simeq 1.96$ V

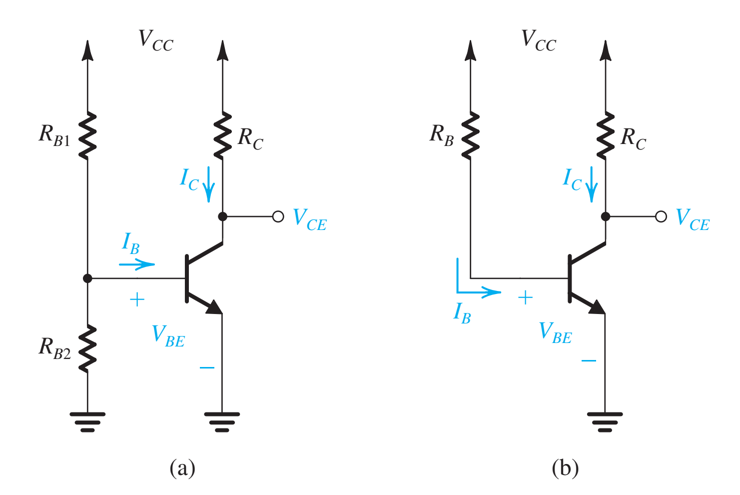


Figure 7.51 Two obvious schemes for biasing the BJT: (**a**) by fixing V_{BE}; (**b**) by fixing I_B. Both result in wide variations in I_C and hence in V_{CE} and therefore are considered to be "bad." Neither scheme is recommended.

7.4.2 The BJT Case

Before presenting the "good" biasing schemes, we should point out why two obvious arrangements are *not* good. First, attempting to bias the BJT by fixing the voltage V_{BE} by, for instance, using a voltage divider across the power supply V_{CC}, as shown in Fig. 7.51(a), is not a viable approach: The very sharp exponential relationship $i_C - v_{BE}$ means that any small and inevitable differences in V_{BE} from the desired value will result in large differences in I_C and in V_{CE}. Second, biasing the BJT by establishing a constant current in the base, as shown in Fig. 7.51(b), where $I_B \simeq (V_{CC} - 0.7)/R_B$, is also not a recommended approach. Here the typically large variations in the value of β among units of the same device type will result in correspondingly large variations in I_C and hence in V_{CE}.

The Classical Discrete-Circuit Bias Arrangement Figure 7.52(a) shows the arrangement most commonly used for biasing a discrete-circuit transistor amplifier if only a single power supply is available. The technique consists of supplying the base of the transistor with a fraction of the supply voltage V_{CC} through the voltage divider R_1, R_2. In addition, a resistor R_E is connected to the emitter. This circuit is very similar to one we used for the MOSFET [Fig. 7.48(c)]. Here, however, the design must take into account the finite base current.

Figure 7.52(b) shows the same circuit with the voltage-divider network replaced by its Thévenin equivalent,

$$V_{BB} = \frac{R_2}{R_1 + R_2} V_{CC} \tag{7.141}$$

$$R_B = \frac{R_1 R_2}{R_1 + R_2} \tag{7.142}$$

The current I_E can be determined by writing a Kirchhoff loop equation for the base–emitter–ground loop, labeled L, and substituting $I_B = I_E/(\beta + 1)$:

$$I_E = \frac{V_{BB} - V_{BE}}{R_E + R_B/(\beta + 1)} \tag{7.143}$$

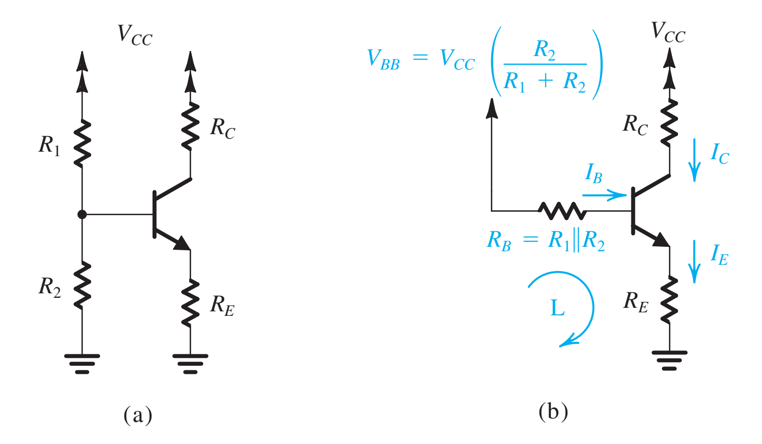


Figure 7.52 Classical biasing for BJTs using a single power supply: (**a**) circuit; (**b**) circuit with the voltage divider supplying the base replaced with its Thévenin equivalent.

To make I_E insensitive to temperature and β variation,[8] we design the circuit to satisfy the following two constraints:

$$V_{BB} \gg V_{BE} \tag{7.144}$$

$$R_E \gg \frac{R_B}{\beta + 1} \tag{7.145}$$

Condition (7.144) ensures that small variations in V_{BE} ($\simeq 0.7$ V) will be swamped by the much larger V_{BB}. There is a limit, however, on how large V_{BB} can be: For a given value of the supply voltage V_{CC}, the higher the value we use for V_{BB}, the lower will be the sum of voltages across R_C and the collector–base junction (V_{CB}). On the other hand, we want the voltage across R_C to be large in order to obtain high voltage gain and large signal swing (before transistor cutoff). We also want V_{CB} (or V_{CE}) to be large, to provide a large signal swing (before transistor saturation). Thus, as is the case in any design, we have a set of conflicting requirements, and the solution must be a trade-off. As a rule of thumb, one designs for V_{BB} about $\frac{1}{3}V_{CC}$, V_{CB} (or V_{CE}) about $\frac{1}{3}V_{CC}$, and $I_C R_C$ about $\frac{1}{3}V_{CC}$.

Condition (7.145) makes I_E insensitive to variations in β and could be satisfied by selecting R_B small. This in turn is achieved by using low values for R_1 and R_2. Lower values for R_1 and R_2, however, will mean a higher current drain from the power supply, and will result in a lowering of the input resistance of the amplifier (if the input signal is coupled to the base),[9] which is the trade-off involved in this part of the design. It should be noted that condition (7.145) means that we want to make the base voltage independent of the value of β and determined solely by the voltage divider. This will obviously be satisfied if the current in the divider is made much larger than the base current. Typically one selects R_1 and R_2 such that their current is in the range of I_E to $0.1I_E$.

Further insight regarding the mechanism by which the bias arrangement of Fig. 7.52(a) stabilizes the dc emitter (and hence collector) current is obtained by considering the feedback

[8]Bias design seeks to stabilize either I_E or I_C since $I_C = \alpha I_E$ and α varies very little. That is, a stable I_E will result in an equally stable I_C, and vice versa.

[9]If the input signal is coupled to the transistor base, the two bias resistances R_1 and R_2 effectively appear in parallel between the base and ground. Thus, low values for R_1 and R_2 will result in lowering R_{in}.

action provided by R_E. Consider that for some reason the emitter current increases. The voltage drop across R_E, and hence V_E, will increase correspondingly. Now, if the base voltage is determined primarily by the voltage divider R_1, R_2, which is the case if R_B is small, it will remain constant, and the increase in V_E will result in a corresponding decrease in V_{BE}. This in turn reduces the collector (and emitter) current, a change opposite to that originally assumed. Thus R_E provides a *negative feedback* action that stabilizes the bias current. This should remind the reader of the resistance R_e that we included in the emitter lead of the CE amplifier in Section 7.3.4. Also, the feedback action of R_E in the circuit of Fig. 7.52(a) is similar to the feedback action of R_S in the circuit of Fig. 7.48(c). We shall study negative feedback formally in Chapter 11.

Example 7.12

We wish to design the bias network of the amplifier in Fig. 7.52 to establish a current $I_E = 1$ mA using a power supply $V_{CC} = +12$ V. The transistor is specified to have a nominal β value of 100.

Solution

We shall follow the rule of thumb mentioned above and allocate one-third of the supply voltage to the voltage drop across R_2 and another one-third to the voltage drop across R_C, leaving one-third for possible negative signal swing at the collector. Thus,

$$V_B = +4 \text{ V}$$
$$V_E = 4 - V_{BE} \simeq 3.3 \text{ V}$$

and R_E is determined from

$$R_E = \frac{V_E}{I_E} = \frac{3.3}{1} = 3.3 \text{ k}\Omega$$

From the discussion above we select a voltage-divider current of $0.1I_E = 0.1 \times 1 = 0.1$ mA. Neglecting the base current, we find

$$R_1 + R_2 = \frac{12}{0.1} = 120 \text{ k}\Omega$$

and

$$\frac{R_2}{R_1 + R_2} V_{CC} = 4 \text{ V}$$

Thus $R_2 = 40 \text{ k}\Omega$ and $R_1 = 80 \text{ k}\Omega$.

At this point, it is desirable to find a more accurate estimate for I_E, taking into account the nonzero base current. Using Eq. (7.143),

$$I_E = \frac{4 - 0.7}{3.3(\text{ k}\Omega) + \frac{(80 \| 40)(\text{ k}\Omega)}{101}} = 0.93 \text{ mA}$$

Example 7.12 *continued*

This is quite a bit lower than 1 mA, the value we are aiming for. It is easy to see from the above equation that a simple way to restore I_E to its nominal value would be to reduce R_E from 3.3 kΩ by the magnitude of the second term in the denominator (0.267 kΩ). Thus a more suitable value for R_E in this case would be $R_E = 3$ kΩ, which results in $I_E = 1.01$ mA $\simeq 1$ mA.[10]

It should be noted that if we are willing to draw a higher current from the power supply and to accept a lower input resistance for the amplifier, then we may use a voltage-divider current equal, say, to I_E (i.e., 1 mA), resulting in $R_1 = 8$ kΩ and $R_2 = 4$ kΩ. We shall refer to the circuit using these latter values as design 2, for which the actual value of I_E using the initial value of R_E of 3.3 kΩ will be

$$I_E = \frac{4 - 0.7}{3.3 + 0.027} = 0.99 \simeq 1 \text{ mA}$$

In this case, design 2, we need not change the value of R_E.

Finally, the value of R_C can be determined from

$$R_C = \frac{12 - V_C}{I_C}$$

Substituting $I_C = \alpha I_E = 0.99 \times 1 = 0.99$ mA $\simeq 1$ mA results, for both designs, in

$$R_C = \frac{12 - 8}{1} = 4 \text{ k}\Omega$$

EXERCISE

7.34 For design 1 in Example 7.12, calculate the expected range of I_E if the transistor used has β in the range of 50 to 150. Express the range of I_E as a percentage of the nominal value ($I_E \simeq 1$ mA) obtained for $\beta = 100$. Repeat for design 2.

Ans. For design 1: 0.94 mA to 1.04 mA, a 10% range; for design 2: 0.984 mA to 0.995 mA, a 1.1% range.

A Two-Power-Supply Version of the Classical Bias Arrangement A somewhat simpler bias arrangement is possible if two power supplies are available, as shown in Fig. 7.53.

[10] Although reducing R_E restores I_E to the design value of 1 mA, it does not solve the problem of the dependence of the value of I_E on β. See Exercise 7.34.

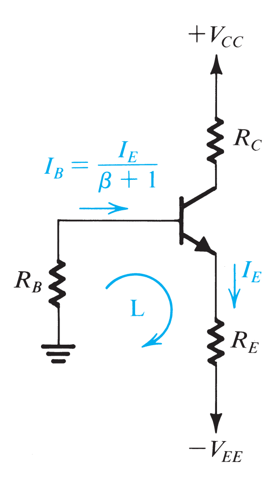


Figure 7.53 Biasing the BJT using two power supplies. Resistor R_B is needed only if the signal is to be capacitively coupled to the base. Otherwise, the base can be connected directly to ground, or to a grounded signal source, resulting in almost total β-independence of the bias current.

Writing a loop equation for the loop labeled L gives

$$I_E = \frac{V_{EE} - V_{BE}}{R_E + R_B/(\beta + 1)} \tag{7.146}$$

This equation is identical to Eq. (7.143) except for V_{EE} replacing V_{BB}. Thus the two constraints of Eqs. (7.144) and (7.145) apply here as well. Note that if the transistor is to be used with the base grounded (i.e., in the common-base configuration), then R_B can be eliminated altogether. On the other hand, if the input signal is to be coupled to the base, then R_B is needed. We shall study complete circuits of the various BJT amplifier configurations in Section 7.5. Finally, observe that the circuit in Fig. 7.53 is the counterpart of the MOS circuit in Fig. 7.48(e).

EXERCISE

D7.35 The bias arrangement of Fig. 7.53 is to be used for a common-base amplifier. Design the circuit to establish a dc emitter current of 1 mA and provide the highest possible voltage gain while allowing for a signal swing at the collector of ±2 V. Use +10-V and –5-V power supplies.
Ans. $R_B = 0$; $R_E = 4.3\ \text{k}\Omega$; $R_C = 8.4\ \text{k}\Omega$

Biasing Using a Collector-to-Base Feedback Resistor In the BJT case, there is a counterpart to the MOSFET circuit of Fig. 7.50. Figure 7.54(a) shows such a simple but effective biasing arrangement that is suitable for common-emitter amplifiers. The circuit employs a resistor R_B connected between the collector and the base. Resistor R_B provides negative feedback, which helps to stabilize the bias point of the BJT.

Figure 7.54 (**a**) A common-emitter transistor amplifier biased by a feedback resistor R_B. (**b**) Analysis of the circuit in (**a**).

Analysis of the circuit is shown in Fig. 7.54(b), from which we can write

$$V_{CC} = I_E R_C + I_B R_B + V_{BE}$$
$$= I_E R_C + \frac{I_E}{\beta + 1} R_B + V_{BE}$$

Thus the emitter bias current is given by

$$I_E = \frac{V_{CC} - V_{BE}}{R_C + R_B/(\beta + 1)} \tag{7.147}$$

It is interesting to note that this equation is identical to Eq. (7.143), which governs the operation of the traditional bias circuit, except that V_{CC} replaces V_{BB} and R_C replaces R_E. It follows that to obtain a value of I_E that is insensitive to variation of β, we select $R_B/(\beta + 1) \ll R_C$. Note, however, that the value of R_B determines the allowable negative signal swing at the collector since

$$V_{CB} = I_B R_B = I_E \frac{R_B}{\beta + 1} \tag{7.148}$$

EXERCISE

D7.36 Design the circuit of Fig. 7.54 to obtain a dc emitter current of 1 mA, maximum gain, and a ±2-V signal swing at the collector; that is, design for $V_{CE} = +2.3$ V. Let $V_{CC} = 10$ V and $\beta = 100$.
Ans. $R_B = 162$ kΩ; $R_C = 7.7$ kΩ. Note that if standard 5% resistor values are used (Appendix J), we select $R_B = 160$ kΩ and $R_C = 7.5$ kΩ. This results in $I_E = 1.02$ mA and $V_C = +2.3$ V.

7.5 Discrete-Circuit Amplifiers

With our study of transistor amplifier basics complete, we now put everything together by presenting practical circuits for discrete-circuit amplifiers. These circuits, which utilize the amplifier configurations studied in Section 7.3 and the biasing methods of Section 7.4, can be assembled using off-the-shelf discrete transistors, resistors, and capacitors. Though practical and carefully selected to illustrate some important points, the circuits presented in this section should be regarded as examples of discrete-circuit transistor amplifiers. Indeed, there is a great variety of such circuits, a number of which are explored in the end-of-chapter problems.

As mentioned earlier, the vast majority of discrete-circuit amplifiers utilize BJTs. This is reflected in this section where all the circuits presented except for one utilize BJTs. Of course, if desired, one can utilize MOSFETs in the same amplifier configurations presented here. Also, the MOSFET is the device of choice in the design of integrated-circuit (IC) amplifiers. We begin our study of IC amplifiers in Chapter 8.

As will be seen shortly, the circuits presented in this section utilize large capacitors (in the μF range) to couple the signal source to the input of the amplifier, and to couple the amplifier output signal to a load resistance or to the input of another amplifier stage. As well, a large capacitor is employed to establish a signal ground at the desired terminal of the transistor (e.g., at the emitter of a CE amplifier). The use of capacitors for these purposes simplifies the design considerably: Since capacitors block dc, one is able to first carry out the dc bias design and then connect the signal source and load to the amplifier without disturbing the dc design. There amplifiers are therefore known as **capacitively coupled amplifiers.**

7.5.1 A Common-Source (CS) Amplifier

As mentioned in Section 7.3, the common-source (CS) configuration is the most widely used of all MOSFET amplifier circuits. A common-source amplifier realized using the bias circuit of Fig. 7.48(c) is shown in Fig. 7.55(a). Observe that to establish a **signal ground**, or an **ac ground** as it is sometimes called, at the source, we have connected a large capacitor, C_S, between the source and ground. This capacitor, usually in the microfarad range, is required to provide a very small impedance (ideally, zero impedance—i.e., in effect, a short circuit) at all signal frequencies of interest. In this way, the signal current passes through C_S to ground and thus *bypasses* the resistance R_S; hence, C_S is called a **bypass capacitor**. Obviously, the lower the signal frequency, the less effective the bypass capacitor becomes. This issue will be studied in Section 10.1. For our purposes here we shall assume that C_S is acting as a perfect short circuit and thus is establishing a zero signal voltage at the MOSFET source.

To prevent disturbances to the dc bias current and voltages, the signal to be amplified, shown as voltage source v_{sig} with an internal resistance R_{sig}, is connected to the gate through a large capacitor C_{C1}. Capacitor C_{C1}, known as a **coupling capacitor**, is required to act as a perfect short circuit at all signal frequencies of interest while blocking dc. Here again, we note that as the signal frequency is lowered, the impedance of C_{C1} (i.e., $1/j\omega C_{C1}$) will increase and its effectiveness as a coupling capacitor will be correspondingly reduced. This problem, too, will be considered in Section 10.1 in connection with the dependence of the amplifier

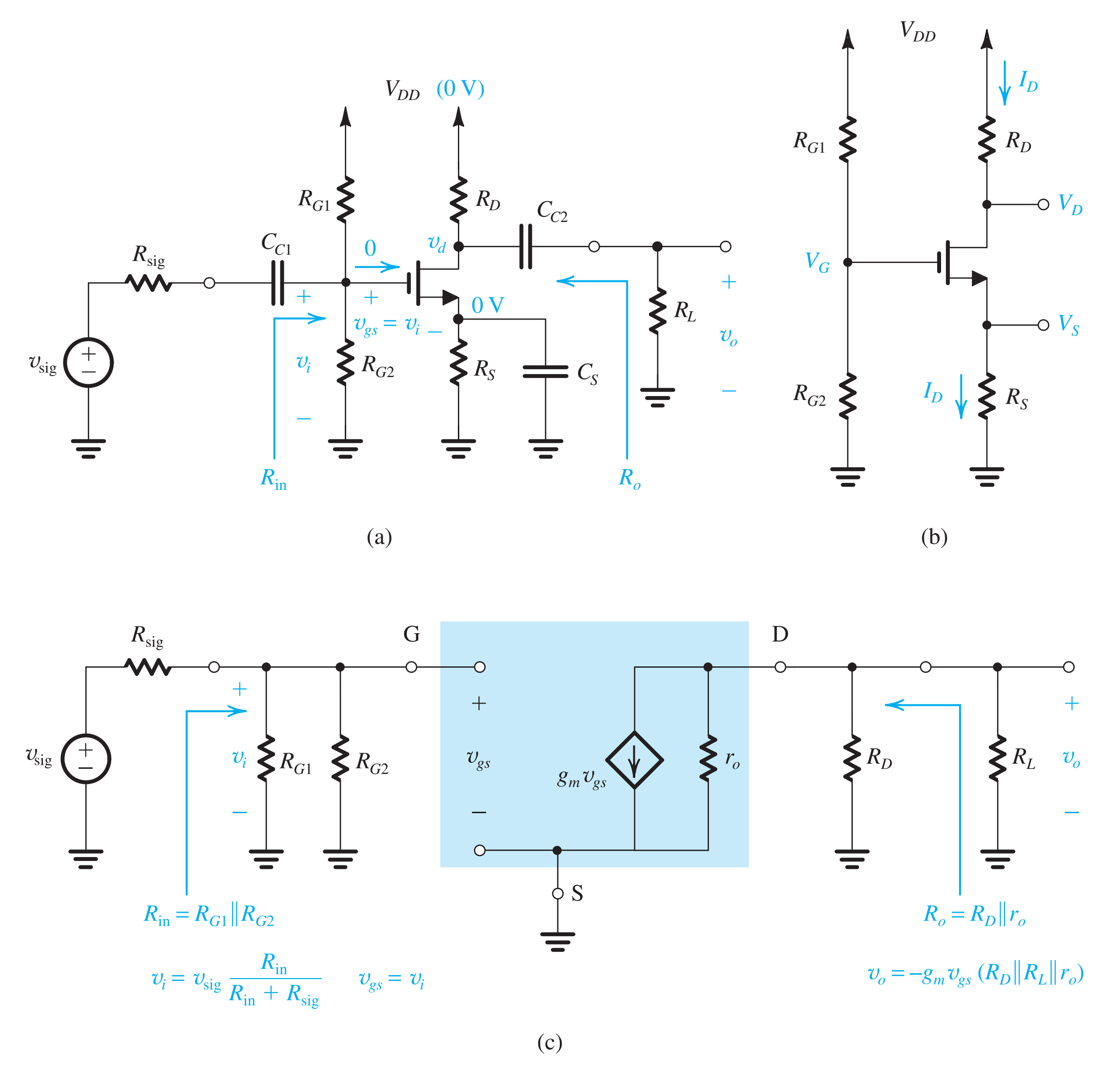


Figure 7.55 (**a**) A common-source amplifier using the classical biasing arrangement of Fig. 7.48(c). (**b**) Circuit for determining the bias point. (**c**) Equivalent circuit and analysis.

operation on frequency. For our purposes here we shall assume that C_{C1} is acting as a perfect short circuit as far as the signal is concerned.

The voltage signal resulting at the drain is coupled to the load resistance R_L via another coupling capacitor C_{C2}. We shall assume that C_{C2} acts as a perfect short circuit at all signal frequencies of interest and thus that the output voltage $v_o = v_d$. Note that R_L can be either an actual load resistor, to which the amplifier is required to provide its output voltage signal, or it can be the input resistance of another amplifier stage in cases where more than one stage of amplification is needed. (We will study multistage amplifiers in Chapter 9).

Since a capacitor behaves as an open circuit at dc, the circuit for performing the dc bias design and analysis is obtained by open-circuiting all capacitors. The resulting circuit is shown in Fig. 7.55(b) and can be designed as discussed in Section 7.4.1.

To determine the terminal characteristics of the CS amplifier of Fig. 7.55(a)—that is, its input resistance, voltage gain, and output resistance—we replace the MOSFET with its hybrid-π small-signal model, replace V_{DD} with a signal ground, and replace all coupling and bypass capacitors with short circuits. The result is the circuit in Fig. 7.55(c). Analysis is straightforward and is shown on the figure, thus

$$R_{\text{in}} = R_{G1} \| R_{G2} \tag{7.149}$$

which shows that to keep R_{in} high, large values should be used for R_{G1} and R_{G2}, usually in the megohm range. The overall voltage gain G_v is

$$G_v = -\frac{R_{\text{in}}}{R_{\text{in}} + R_{\text{sig}}} g_m (R_D \| R_L \| r_o) \tag{7.150}$$

Observe that we have taken r_o into account, simply because it is easy to do so. Its effect, however, is usually small (this is not the case for IC amplifiers, as will be explained in Chapter 8). Finally, to encourage the reader to do the small-signal analysis directly on the original circuit diagram, with the MOSFET model used implicitly, we show some of the analysis on the circuit of Fig. 7.55(a).

EXERCISES

D7.37 Design the bias circuit in Fig.7.55(b) for the CS amplifier of Fig. 7.55(a). Assume the MOSFET is specified to have $V_t = 1$ V, $k_n = 4$ mA/V^2, and $V_A = 100$ V. Neglecting the Early effect, design for $I_D = 0.5$ mA, $V_S = 3.5$ V, and $V_D = 6$ V using a power-supply $V_{DD} = 15$ V. Specify the values of R_S and R_D. If a current of 2 μA is used in the voltage divider, specify the values of R_{G1} and R_{G2}. Give the values of the MOSFET parameters g_m and r_o at the bias point.
Ans. $R_S = 7$ kΩ; $R_D = 18$ kΩ; $R_{G1} = 5$ MΩ; $R_{G2} = 2.5$ MΩ; $g_m = 2$ mA/V; $r_o = 200$ kΩ

7.38 For the CS amplifier of Fig. 7.55(a) use the design obtained in Exercise 7.37 to determine R_{in}, R_o, and the overall voltage gain G_v when $R_{\text{sig}} = 100$ kΩ and $R_L = 20$ kΩ.
Ans. 1.67 MΩ; 16.5 kΩ; −17.1 V/V

D7.39 As discussed in Section 7.3, beneficial effects can be realized by having an unbypassed resistance R_s in the source lead of the CS amplifier. This can be implemented in the circuit of Fig. 7.55(a) by splitting the resistance R_S into two resistances: R_s, which is left unbypassed, and $(R_S - R_s)$, across which the bypass capacitor C_S is connected. Now, if in order to improve linearity of the amplifier in Exercises 7.37 and 7.38, v_{gs} is to be reduced to half its value, what value should R_s have? What would the amplifier gain G_v become? Recall that when R_s is included it becomes difficult to include r_o in the analysis, so neglect it.
Ans. $R_s = 500$ Ω; $G_v = -8.9$ V/V

7.5.2 A Common-Emitter Amplifier

The common-emitter (CE) amplifier is the most widely used of all BJT amplifier configurations. Figure 7.56(a) shows a CE amplifier utilizing the classical biasing arrangement of Fig. 7.48(c), the design of which was considered in Section 7.4. The CE circuit in Fig. 7.54(a) is the BJT counterpart of the CS amplifier of Fig. 7.55(a). It utilizes coupling capacitors C_{C1} and C_{C2} and bypass capacitor C_E. Here we assume that these capacitors, while blocking dc, behave as perfect short circuits at all signal frequencies of interest.

To determine the characteristic parameters of the CE amplifier, we replace the BJT with its hybrid-π model, replace V_{CC} with a short circuit to ground, and replace the coupling and bypass capacitor with short circuits. The resulting small-signal equivalent circuit of the CE amplifier is shown in Fig. 7.56(b). The analysis is straightforward and is given in the

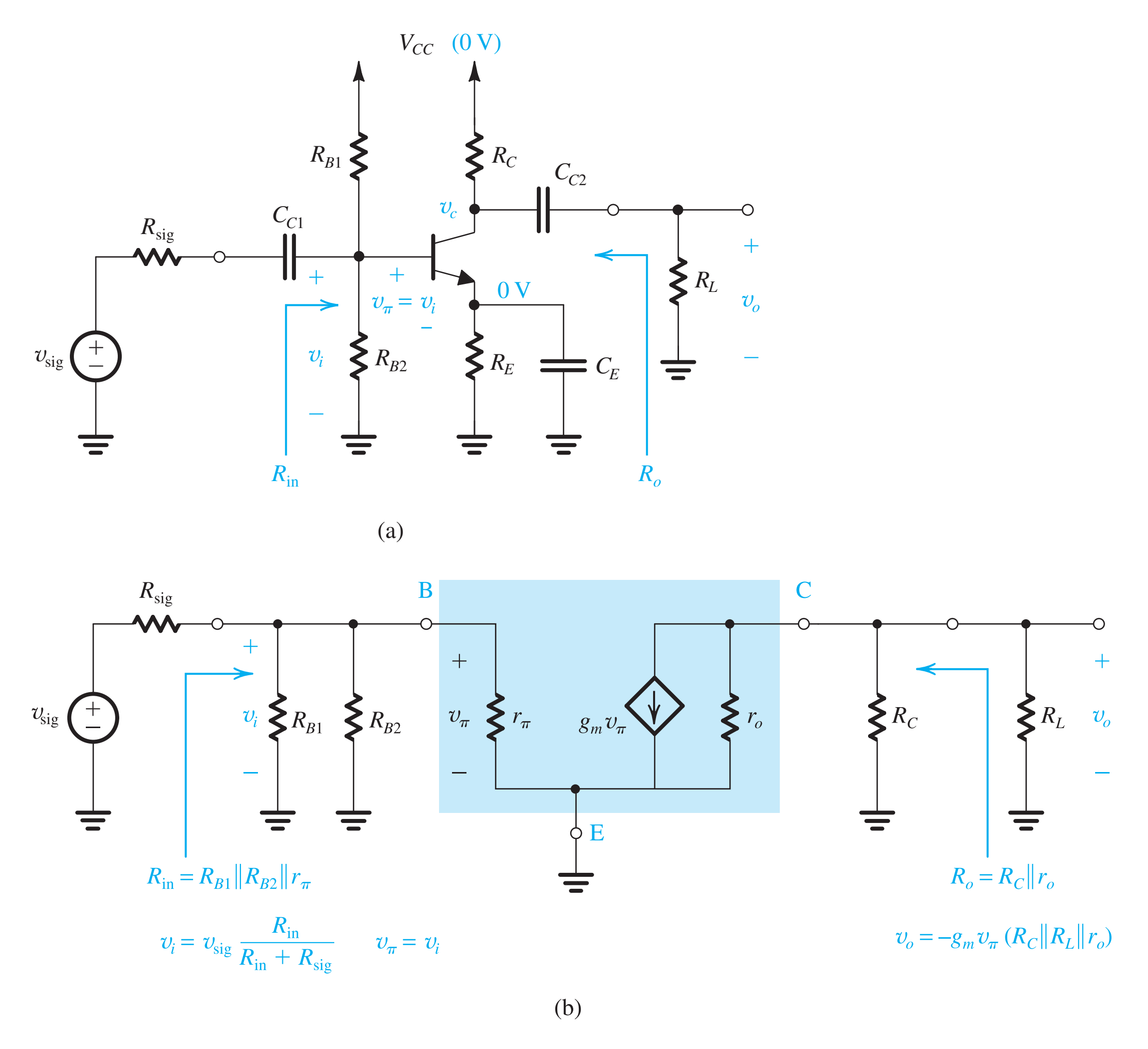


Figure 7.56 (**a**) A common-emitter amplifier using the classical biasing arrangement of Fig. 7.52(a). (**b**) Equivalent circuit and analysis.

figure, thus

$$R_{in} = R_{B1} \| R_{B2} \| r_\pi \quad (7.151)$$

which indicates that to keep R_{in} relatively high, R_{B1} and R_{B2} should be selected large (typically in the range of tens or hundreds of kilohms). This requirement conflicts with the need to keep R_{B1} and R_{B2} low so as to minimize the dependence of the dc current I_C on the transistor β. We discussed this design trade-off in Section 7.4.

The voltage gain G_v is given by

$$G_v = -\frac{R_{in}}{R_{in} + R_{sig}} g_m (R_C \| R_L \| r_o) \quad (7.152)$$

Note that we have taken r_o into account because it is easy to do so. However, as already mentioned, the effect of this parameter on discrete-circuit amplifier performance is usually small.

EXERCISES

D7.40 Design the bias circuit of the CE amplifier of Fig. 7.56(a) to obtain $I_E = 0.5$ mA and $V_C = +6$ V. Design for a dc voltage at the base of 5 V and a current through R_{B2} of 50 μA. Let $V_{CC} = +15$ V, $\beta = 100$, and $V_{BE} \simeq 0.7$ V. Specify the values of R_{B1}, R_{B2}, R_E, and R_C. Also give the values of the BJT small-signal parameters g_m, r_π, and r_o at the bias point. (For the calculation of r_o, let $V_A = 100$ V.)
Ans. $R_{B1} = 182$ kΩ; $R_{B2} = 100$ kΩ; $R_E = 8.6$ kΩ; $R_C = 18$ kΩ; $g_m = 20$ mA/V, $r_\pi = 5$ kΩ, $r_o = 200$ kΩ

7.41 For the amplifier designed in Exercise 7.40, find R_{in}, R_o, and G_v when $R_{sig} = 10$ kΩ and $R_L = 20$ kΩ.
Ans. $R_{in} = 4.64$ kΩ; $R_o = 16.51$ kΩ; $G_v = -57.3$ V/V

7.5.3 A Common-Emitter Amplifier with an Emitter Resistance R_e

As discussed in Section 7.3.4, it is beneficial to include a small resistance in the transistor emitter lead. This can be implemented in the circuit of Fig. 7.56(a) by splitting the emitter bias resistance R_E into two components: an unbypassed resistance R_e, and a resistance $(R_E - R_e)$ across which the bypass capacitor C_E is connected. The resulting circuit is shown in Fig. 7.57(a) and its small-signal model is shown in Fig. 7.57(b). In the latter we utilize the T model of the BJT because it results in much simpler analysis (recall that this is always the case when a resistance is connected in series with the emitter). Also note that we have not included r_o, for doing so would complicate the analysis significantly. This burden would not be justified given that r_o has little effect on the performance of discrete-circuit amplifiers.

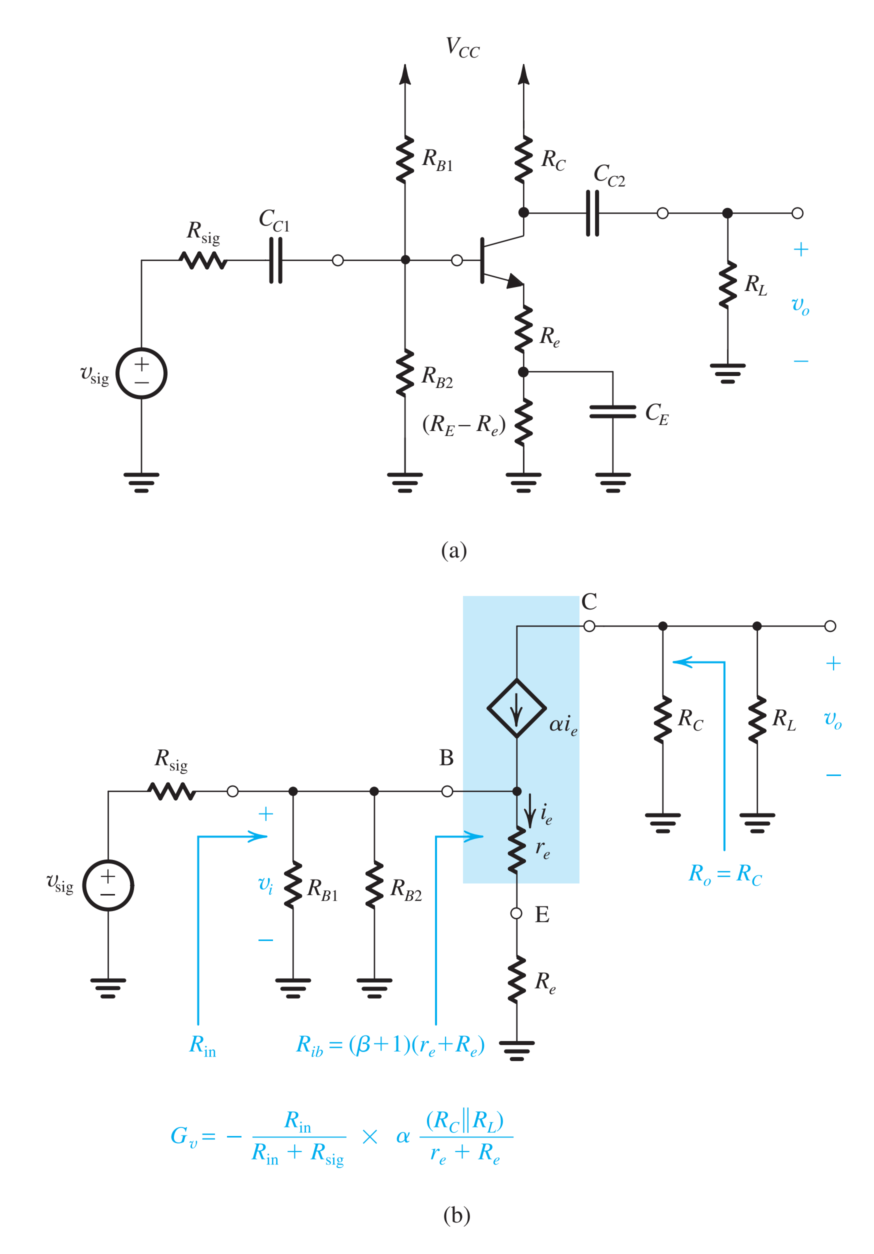


Figure 7.57 (**a**) A common-emitter amplifier with an unbiased emitter resistance R_e. (**b**) The amplifier small-signal model and analysis.

Analysis of the circuit in Fig. 7.57(b) is straightforward and is shown in the figure. Thus,

$$\begin{aligned} R_{in} &= R_{B1} \| R_{B2} \| (\beta+1)(r_e + R_e) \\ &= R_{B1} \| R_{B2} \| [r_\pi + (\beta+1)R_e] \end{aligned} \tag{7.153}$$

from which we note that including R_e increases R_{in} because it increases the input resistance looking into the base by adding a component $(\beta + 1)R_e$ to r_π. The overall voltage gain G_v is

$$G_v = -\frac{R_{in}}{R_{in}+R_{sig}} \times \alpha \frac{\text{Total resistance in collector}}{\text{Total resistance in emitter}}$$

$$= -\alpha \frac{R_{in}}{R_{in}+R_{sig}} \frac{R_C \parallel R_L}{r_e + R_e} \tag{7.154}$$

EXERCISE

7.42 For the amplifier designed in Exercise 7.40 and analyzed in Exercise 7.41, let it be required to raise R_{in} to 10 kΩ. What is the required value of R_e, and what does the overall voltage gain G_v become?
Ans. $R_e = 67.7\ \Omega$; $G_v = -39.8$ V/V

7.5.4 A Common-Base (CB) Amplifier

Figure 7.58(a) shows a CB amplifier designed using the biasing arrangement of Fig. 7.53. Note that the availability of two power supplies, V_{CC} and $-V_{EE}$, enables us to connect the base directly to ground, obviating the need for a large bypass capacitor to establish a signal ground at the base.

The small-signal equivalent circuit of the CB amplifier is shown in Fig 7.58(b). As expected, we have utilized the T model of the BJT and have not included r_o. Including r_o would complicate the analysis significantly without making much difference to the results in the case of discrete-circuit amplifiers. From the circuit in Fig. 7.58(b) we find

$$R_{in} = r_e \parallel R_E \simeq r_e \simeq 1/g_m$$

which as expected can be very small, causing v_i to be a small fraction of v_{sig},

$$v_i = v_{sig} \frac{R_{in}}{R_{in}+R_{sig}}$$

Now,

$$i_e = -\frac{v_i}{r_e}$$

and

$$v_o = -\alpha i_e (R_C \parallel R_L)$$

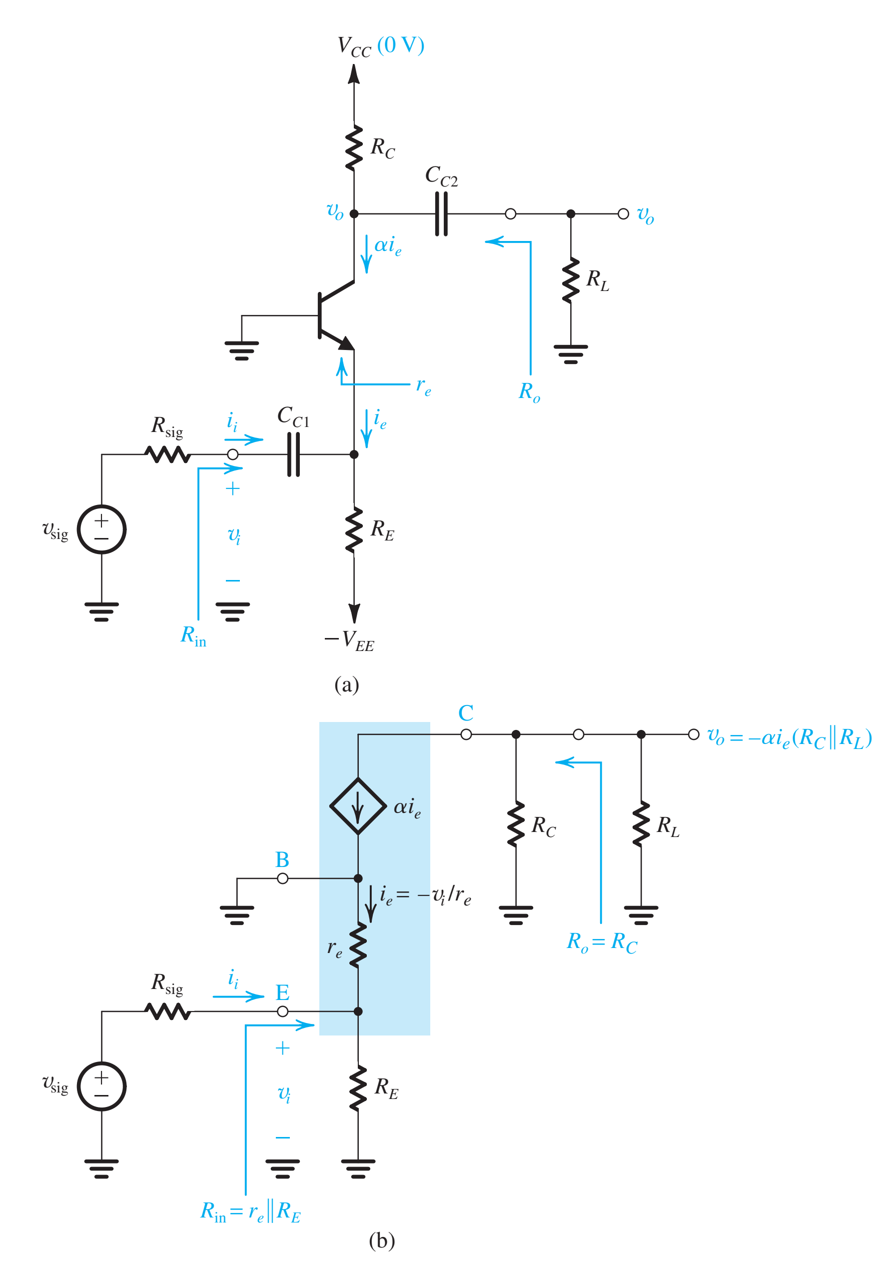


Figure 7.58 (**a**) A common-base amplifier using the structure of Fig. 7.53 with R_B omitted (since the base is grounded). (**b**) Equivalent circuit obtained by replacing the transistor with its T model.

Thus, the overall voltage gain is given by

$$G_v = \alpha \frac{R_{in}}{R_{in} + R_{sig}} \frac{R_C \| R_L}{r_e} = \frac{R_{in}}{R_{in} + R_{sig}} g_m (R_C \| R_L) \tag{7.155}$$

EXERCISE

D7.43 Design the CB amplifier of Fig. 7.58(a) to provide an input resistance R_{in} that matches the source resistance of a cable with a characteristic resistance of 50 Ω. Assume that $R_E \gg r_e$. The available power supplies are ±5 V and $R_L = 8$ kΩ. Design for a dc collector voltage $V_C = +1$ V. Specify the values of R_C and R_E. What overall voltage gain is obtained? If v_{sig} is a sine wave with a peak amplitude of 10 mV, what is the peak amplitude of the output voltage? Let $\alpha \simeq 1$.
Ans. $R_C = 8$ kΩ; $R_E = 8.6$ kΩ; 40 V/V; 0.4 V

7.5.5 An Emitter Follower

Figure 7.59(a) shows an emitter follower designed using the bias arrangement of Fig. 7.53 and two power supplies, V_{CC} and $-V_{EE}$. The bias resistance R_B affects the input resistance of the follower and should be chosen as large as possible while limiting the dc voltage drop across it to a small fraction of V_{EE}; otherwise the dependence of the bias current I_C on β can become unacceptably large.

Figure 7.59(b) shows the small-signal equivalent circuit of the emitter follower. Here, as expected, we have replaced the BJT with its T model and included r_o (since this can be done very simply). The input resistance of the emitter follower can be seen to be

$$R_{in} = R_B \| R_{ib} \tag{7.156}$$

where R_{ib}, the input resistance looking into the base, can be obtained by using the resistance-reflection rule. Toward that end, note that r_o appears in parallel with R_E and R_L (which is why it can be easily taken into account). Thus,

$$R_{ib} = (\beta + 1)[r_e + (R_E \| r_o \| R_L)] \tag{7.157}$$

The overall voltage gain can be determined by tracking the signal transmission from source to load,

$$v_i = v_{sig} \frac{R_{in}}{R_{in} + R_{sig}} \tag{7.158}$$

and

$$v_o = v_i \frac{R_E \| r_o \| R_L}{r_e + (R_E \| r_o \| R_L)} \tag{7.159}$$

Thus,

$$G_v \equiv \frac{v_o}{v_{sig}} = \frac{R_{in}}{R_{in} + R_{sig}} \frac{(R_E \| r_o \| R_L)}{r_e + (R_E \| r_o \| R_L)} \tag{7.160}$$

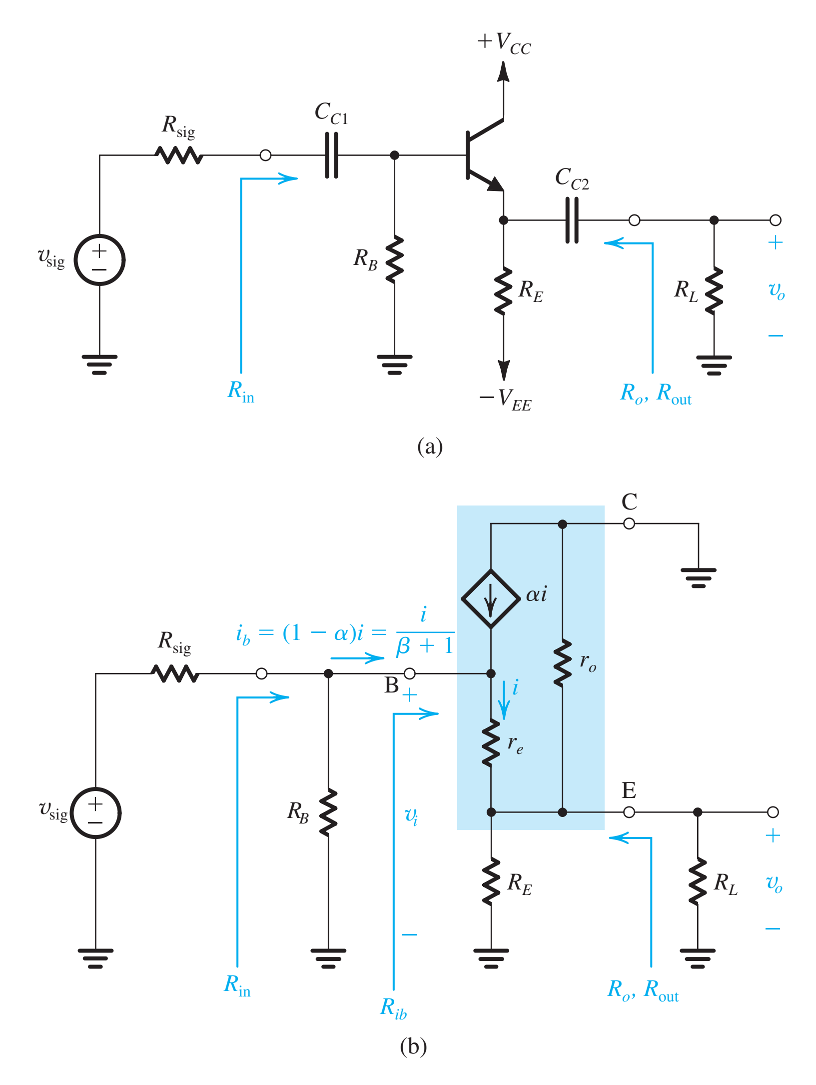


Figure 7.59 (**a**) An emitter-follower circuit. (**b**) Small-signal equivalent circuit of the emitter follower with the transistor replaced by its T model. Note that r_o is included because it is easy to do so. Normally, its effect on performance is small.

Finally, the output resistance R_{out} can be obtained by short-circuiting v_{sig} and looking back into the output terminal, excluding R_L, as

$$R_{out} = r_o \parallel R_E \parallel \left[r_e + \frac{R_B \parallel R_{sig}}{\beta + 1} \right] \tag{7.161}$$

Note that we have used the inverse resistance-reflection rule, namely, dividing the total resistance in the base, $(R_B \parallel R_{sig})$, by $(\beta + 1)$.

EXERCISE

D7.44 Design the emitter follower of Fig. 7.59(a) to operate at a dc emitter current $I_E = 1$ mA. Allow a dc voltage drop across R_B of 1 V. The available power supplies are ±5 V, $\beta = 100$, $V_{BE} = 0.7$ V, and $V_A = 100$ V. Specify the values required for R_B and R_E. Now if $R_{sig} = 50$ kΩ and $R_L = 1$ kΩ, find R_{in}, v_i/v_{sig}, v_o/v_i, G_v, and R_{out}. (*Note*: In performing the bias design, neglect the Early effect.)

Ans. $R_B = 100$ kΩ; $R_E = 3.3$ kΩ; 44.3 kΩ; 0.469 V/V; 0.968 V/V; 0.454 V/V; 320 Ω

7.5.6 The Amplifier Frequency Response

Thus far, we have assumed that the gain of transistor amplifiers is constant independent of the frequency of the input signal. This would imply that transistor amplifiers have infinite bandwidth, which of course is not true. To illustrate, we show in Fig. 7.60 a sketch of the magnitude of the gain of a common-emitter or a CS amplifier such as those shown in Figs. 7.56 and 7.55, respectively, versus frequency. Observe that there is indeed a wide frequency range over which the gain remains almost constant. This obviously is the useful frequency range of

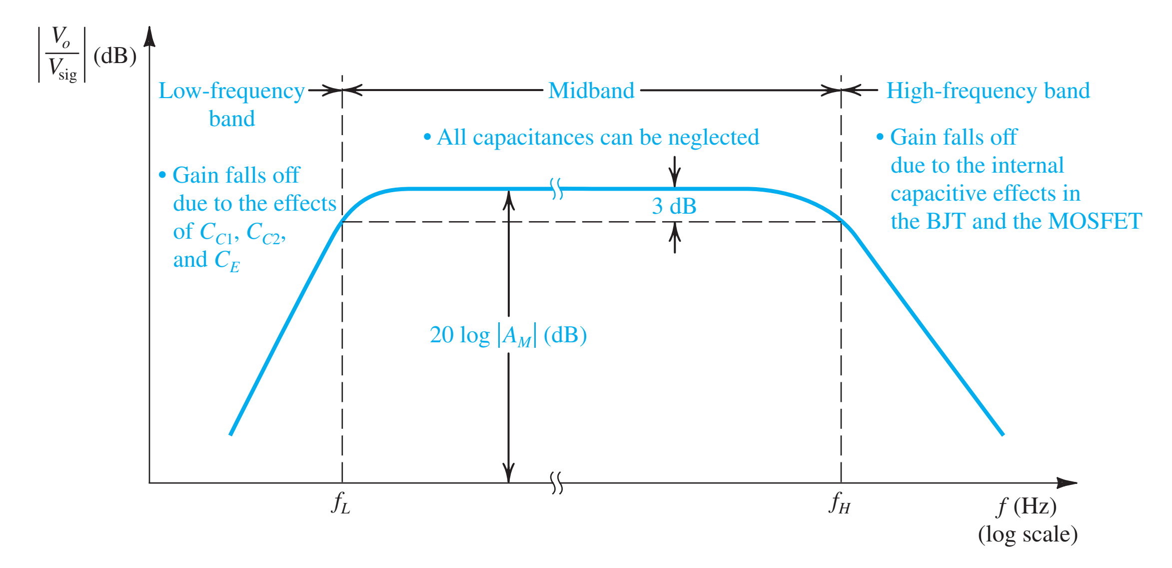


Figure 7.60 Sketch of the magnitude of the gain of a CE (Fig. 7.56) or CS (Fig. 7.55) amplifier versus frequency. The graph delineates the three frequency bands relevant to frequency-response determination.

operation for the particular amplifier. Thus far, we have been assuming that our amplifiers are operating in this frequency band, called the **midband**.

Figure 7.60 indicates that at lower frequencies, the magnitude of amplifier gain falls off. This is because the coupling and bypass capacitors no longer have low impedances. Recall that we assumed that their impedances were small enough to act as short circuits. Although this can be true at midband frequencies, as the frequency of the input signal is lowered, the reactance $1/j\omega C$ of each of these capacitors becomes significant, and it can be shown that this results in the overall voltage gain of the amplifier decreasing.

Figure 7.60 indicates also that the gain of the amplifier falls off at the high-frequency end. This is due to the internal capacitive effects in the BJT and the MOSFET. In Chapter 10 we shall study the internal capacitive effects of both transistor types and will augment their hybrid-π models with capacitances that model these effects.

We will undertake a detailed study of the frequency response of transistor amplifiers in Chapter 10. For the time being, however, it is important for the reader to realize that for every transistor amplifier, there is a finite band over which the gain is almost constant. The boundaries of this useful frequency band, or midband, are the two frequencies f_L and f_H at which the gain drops by a certain number of decibels (usually 3 dB) below its value at midband. As indicated in Fig. 7.60, the amplifier **bandwidth**, or 3-dB bandwidth, is defined as the difference between the lower (f_L) and upper or higher (f_H) 3-dB frequencies:

$$BW = f_H - f_L$$

and since usually $f_L \ll f_H$,

$$BW \simeq f_H$$

A figure of merit for the amplifier is its **gain–bandwidth product**, defined as

$$GB = |A_M| BW$$

where $|A_M|$ is the magnitude of the amplifier gain in the midband. It will be seen in Chapter 10 that in amplifier design it is usually possible to trade off gain for bandwidth. One way to accomplish this, for instance, is by including resistance R_e in the emitter of the CE amplifier.

Summary

- The essence of the use of the MOSFET (the BJT) as an amplifier is that when the transistor is operated in the active region, v_{GS} controls i_D (v_{BE} controls i_C) in the manner of a voltage-controlled current source. When the device is dc biased in the active region, and the signal v_{gs} (v_{be}) is kept small, the operation becomes almost linear, with $i_d = g_m v_{gs}$ ($i_c = g_m v_{be}$).

- The most fundamental parameter in characterizing the small-signal linear operation of a transistor is the transconductance g_m. For a MOSFET, $g_m = \mu_n C_{ox}(W/L)V_{OV} = \sqrt{2\mu_n C_{ox}(W/L)I_D} = 2I_D/V_{OV}$; and for the BJT, $g_m = I_C/V_T$.

- A systematic procedure for the analysis of a transistor amplifier circuit is presented in Table 7.1. Tables 7.2 and 7.3 present the small-signal models for the MOSFET and the BJT, respectively.

- When a resistance is connected in series with the source (or emitter), the T model is the most convenient to use.

- The three basic configurations of MOS and BJT amplifiers are presented in Fig. 7.33. Their characteristic parameter values are provided in Table 7.4 (for the MOS case) and in Table 7.5 (for the BJT case).

- The CS amplifier, which has (ideally) infinite input resistance and a reasonably high gain but a rather high output resistance and a limited high-frequency response (more on the latter point in Chapter 10), is used to obtain most of the gain in a cascade amplifier. Similar remarks apply to the CE amplifier, except that it has a relatively low input resistance ($r_\pi = \beta/g_m$) arising from the finite base current of the BJT (finite β). Its voltage gain, however, can be larger than that of the CS amplifier because of the higher values of g_m obtained with BJTs.

- Adding a resistance R_s in the source of a CS amplifier (a resistance R_e in the emitter of a CE amplifier) can lead to beneficial effects including the following: raising the input resistance of the CE amplifier, increasing linearity, and extending the useful amplifier bandwidth, at the expense of reducing the gain, all by a factor equal to $(1 + g_m R_s)$ [$(1 + g_m R_e)$ for the BJT case].

- The CG (CB) amplifier has a low input resistance and thus, used alone, it has limited and specialized applications. However, its excellent high-frequency response makes it attractive in combination with the CS (CE) amplifier (Chapters 8 and 10).

- The source follower has (ideally) infinite input resistance, a voltage gain lower than but close to unity, and a low output resistance. It is employed as a voltage buffer and as the output stage of a multistage amplifier. Similar remarks apply to the emitter follower except that its input resistance, though large, is finite. Specifically, the emitter follower multiplies the total resistance in the emitter by $(\beta + 1)$ before presenting it to the signal source.

- The resistance-reflection rule is a powerful tool in the analysis of BJT amplifier circuits: All resistances in the emitter circuit including the emitter resistance r_e can be reflected to the base side by multiplying them by $(\beta + 1)$. Conversely, we can reflect all resistances in the base circuit to the emitter side by dividing them by $(\beta + 1)$.

- In the analysis and design of discrete-circuit amplifiers, it is rarely necessary to take the transistor output resistance r_o into account. In some situations, however, r_o can be easily taken into account; specifically in the CS (CE) amplifier and in the source (emitter) follower. In IC amplifiers, r_o must always be taken into account.

- A key step in the design of transistor amplifiers is to bias the transistor to operate at an appropriate point in the active region. A good bias design ensures that the parameters of the operating point (I_D, V_{OV}, and V_{DS} for the MOSFET; I_C and V_{CE} for the BJT) are predictable and stable and do not vary by large amounts when the transistor is replaced by another of the same type. The bias methods studied in this chapter are suited for discrete-circuit amplifiers only because they utilize large coupling and bypass capacitors.

- Discrete-circuit amplifiers predominantly employ BJTs. The opposite is true for IC amplifiers, where the device of choice is the MOSFET.

Computer Simulation Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as allowable signal swing and amplifier nonlinear distortion. Instructions to assist in setting up PSpice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 7.1: Basic Principles

7.1 For the MOS amplifier of Fig. 7.2(a) with $V_{DD} = 5$ V, $V_t = 0.5$ V, $k_n = 10$ mA/V^2, and $R_D = 20$ kΩ, determine the coordinates of the active-region segment (AB) of the VTC [Fig. 7.2(b)].

D 7.2 For the MOS amplifier of Fig. 7.2(a) with $V_{DD} = 5$ V and $k_n = 5$ mA/V^2, it is required to have the end point of the VTC, point B, at $V_{DS} = 0.5$ V. What value of R_D is required? If the transistor is replaced with another having twice the value of the transconductance parameter k_n, what new value of R_D is needed?

D 7.3 It is required to bias the MOS amplifier of Fig. 7.3 at point Q for which $V_{OV} = 0.2$ V and $V_{DS} = 1$ V. Find the required value of R_D when $V_{DD} = 5$ V, $V_t = 0.5$ V, and $k_n = 10$ mA/V^2. Also specify the coordinates of the VTC end point B. What is the small-signal voltage gain of this amplifier? Assuming linear operation, what is the maximum allowable negative signal swing at the output? What is the corresponding peak input signal?

7.4 The MOS amplifier of Fig. 7.4(a), when operated with $V_{DD} = 2$ V, is found to have a maximum small-signal voltage gain magnitude of 14 V/V. Find V_{OV} and V_{DS} for bias point Q at which a voltage gain of −12 V/V is obtained.

7.5 Consider the amplifier of Fig. 7.4(a) for the case $V_{DD} = 5$ V, $R_D = 24$ kΩ, $k_n'(W/L) = 1$ mA/V^2, and $V_t = 1$ V.

(a) Find the coordinates of the two end points of the saturation-region segment of the amplifier transfer characteristic, that is, points A and B on the sketch of Fig. 7.4(b).

(b) If the amplifier is biased to operate with an overdrive voltage V_{OV} of 0.5 V, find the coordinates of the bias point Q on the transfer characteristic. Also, find the value of I_D and of the incremental gain A_v at the bias point.

(c) For the situation in (b), and disregarding the distortion caused by the MOSFET's square-law characteristic, what is the largest amplitude of a sine-wave voltage signal that can be applied at the input while the transistor remains in saturation? What is the amplitude of the output voltage signal that results? What gain value does the combination of these amplitudes imply? By what percentage is this gain value different from the incremental gain value calculated above? Why is there a difference?

7.6 Various measurements are made on an NMOS amplifier for which the drain resistor R_D is 20 kΩ. First, dc measurements show the voltage across the drain resistor, V_{RD}, to be 1.5 V and the gate-to-source bias voltage to be 0.7 V. Then, ac measurements with small signals show the voltage gain to be –10 V/V. What is the value of V_t for this transistor? If the process transconductance parameter k_n' is 200 μA/V^2, what is the MOSFET's W/L?

***7.7** The expression for the incremental voltage gain A_v given in Eq. (7.16) can be written in as

$$A_v = -\frac{2(V_{DD} - V_{DS})}{V_{OV}}$$

where V_{DS} is the bias voltage at the drain. This expression indicates that for given values of V_{DD} and V_{OV}, the gain magnitude can be increased by biasing the transistor at a lower V_{DS}. This, however, reduces the allowable output signal swing in the negative direction. Assuming linear operation around the bias point, show that the largest possible negative output signal peak $\hat{v}_o$ that is achievable while the transistor remains saturated is

$$\hat{v}_o = (V_{DS} - V_{OV}) \Big/ \left(1 + \frac{1}{|A_v|}\right)$$

For $V_{DD} = 5$ V and $V_{OV} = 0.5$ V, provide a table of values for A_v, $\hat{v}_o$, and the corresponding $\hat{v}_i$ for $V_{DS} = 1$ V, 1.5 V, 2 V, and 2.5 V. If $k_n'(W/L) = 1$ mA/V^2, find I_D and R_D for the design for which $V_{DS} = 1$ V.

D *7.8 Design the MOS amplifier of Fig. 7.4(a) to obtain maximum gain while allowing for an output voltage swing of at least ±0.5 V. Let $V_{DD} = 5$ V, and utilize an overdrive

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

voltage of approximately 0.2 V.

(a) Specify V_{DS} at the bias point.
(b) What is the gain achieved? What is the signal amplitude $\hat{v}_{gs}$ that results in the 0.5-V signal amplitude at the output?
(c) If the dc bias current in the drain is to be 100 μA, what value of R_D is needed?
(d) If $k_n' = 200\ \mu A/V^2$, what W/L ratio is required for the MOSFET?

***7.9** Figure P7.9 shows an amplifier in which the load resistor R_D has been replaced with another NMOS transistor Q_2 connected as a two-terminal device. Note that because v_{DG} of Q_2 is zero, it will be operating in saturation at all times, even when $v_I = 0$ and $i_{D2} = i_{D1} = 0$. Note also that the two transistors conduct equal drain currents. Using $i_{D1} = i_{D2}$, show that for the range of v_I over which Q_1 is operating in saturation, that is, for

$$V_{t1} \le v_I \le v_O + V_{t1}$$

the output voltage will be given by

$$v_O = V_{DD} - V_t + \sqrt{\frac{(W/L)_1}{(W/L)_2}}V_t - \sqrt{\frac{(W/L)_1}{(W/L)_2}}v_I$$

where we have assumed $V_{t1} = V_{t2} = V_t$. Thus the circuit functions as a linear amplifier, even for large input signals. For $(W/L)_1 = (50\ \mu m/0.5\ \mu m)$ and $(W/L)_2 = (5\ \mu m/0.5\ \mu m)$, find the voltage gain.

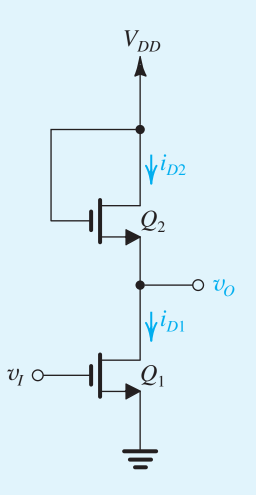


Figure P7.9

7.10 A BJT amplifier circuit such as that in Fig. 7.6 is operated with $V_{CC} = +5$ V and is biased at $V_{CE} = +1$ V. Find the voltage gain, the maximum allowed output negative swing without the transistor entering saturation, and the corresponding maximum input signal permitted.

7.11 For the amplifier circuit in Fig. 7.6 with $V_{CC} = +5$ V and $R_C = 1\ k\Omega$, find V_{CE} and the voltage gain at the following dc collector bias currents: 0.5 mA, 1 mA, 2.5 mA, 4 mA, and 4.5 mA. For each, give the maximum possible positive- and negative-output signal swing as determined by the need to keep the transistor in the active region. Present your results in a table.

D 7.12 Consider the CE amplifier circuit of Fig. 7.6 when operated with a dc supply $V_{CC} = +5$ V. It is required to find the point at which the transistor should be biased; that is, find the value of V_{CE} so that the output sine-wave signal v_{ce} resulting from an input sine-wave signal v_{be} of 5-mV peak amplitude has the maximum possible magnitude. What is the peak amplitude of the output sine wave and the value of the gain obtained? Assume linear operation around the bias point. (*Hint:* To obtain the maximum possible output amplitude for a given input, you need to bias the transistor as close to the edge of saturation as possible without entering saturation at any time, that is, without v_{CE} decreasing below 0.3 V.)

7.13 A designer considers a number of low-voltage BJT amplifier designs utilizing power supplies with voltage V_{CC} of 1.0, 1.5, 2.0, or 3.0 V. For transistors that saturate at $V_{CE} =$ 0.3 V, what is the largest possible voltage gain achievable with each of these supply voltages? If in each case biasing is adjusted so that $V_{CE} = V_{CC}/2$, what gains are achieved? If a negative-going output signal swing of 0.4 V is required, at what V_{CE} should the transistor be biased to obtain maximum gain? What is the gain achieved with each of the supply voltages? (Notice that all of these gains are independent of the value of I_C chosen!)

D *7.14 A BJT amplifier such as that in Fig. 7.6 is to be designed to support relatively undistorted sine-wave output signals of peak amplitudes P volt without the BJT entering saturation or cutoff and to have the largest possible voltage gain, denoted A_v V/V. Show that the minimum supply voltage V_{CC} needed is given by

$$V_{CC} = V_{CEsat} + P + |A_v|V_T$$

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

Also, find V_{CC}, specified to the nearest 0.5 V, for the following situations:

(a) $A_v = -20$ V/V, $P = 0.2$ V
(b) $A_v = -50$ V/V, $P = 0.5$ V
(c) $A_v = -100$ V/V, $P = 0.5$ V
(d) $A_v = -100$ V/V, $P = 1.0$ V
(e) $A_v = -200$ V/V, $P = 1.0$ V
(f) $A_v = -500$ V/V, $P = 1.0$ V
(g) $A_v = -500$ V/V, $P = 2.0$ V

7.15 The transistor in the circuit of Fig. P7.15 is biased at a dc collector current of 0.3 mA. What is the voltage gain? (*Hint:* Use Thévenin's theorem to convert the circuit to the form in Fig. 7.6.)

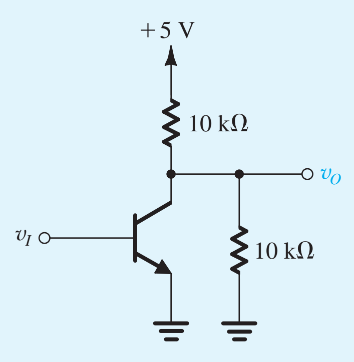


Figure P7.15

7.16 Sketch and label the voltage-transfer characteristics of the *pnp* amplifiers shown in Fig. P7.16.

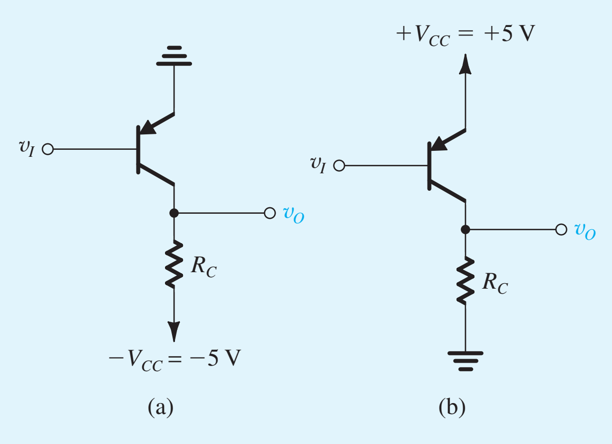


Figure P7.16

***7.17** In deriving the expression for small-signal voltage gain A_v in Eq. (7.21) we neglected the Early effect. Derive this expression including the Early effect by substituting

$$i_C = I_S e^{v_{BE}/V_T}\left(1 + \frac{v_{CE}}{V_A}\right)$$

in Eq. (7.4) and including the factor $(1 + V_{CE}/V_A)$ in Eq. (7.11). Show that the gain expression changes to

$$A_v = \frac{-I_C R_C/V_T}{\left[1 + \dfrac{I_C R_C}{V_A + V_{CE}}\right]} = -\frac{(V_{CC} - V_{CE})/V_T}{\left[1 + \dfrac{V_{CC} - V_{CE}}{V_A + V_{CE}}\right]}$$

For the case $V_{CC} = 5$ V and $V_{CE} = 3$ V, what is the gain without and with the Early effect taken into account? Let $V_A = 100$ V.

7.18 When the amplifier circuit of Fig. 7.6 is biased with a certain V_{BE}, the dc voltage at the collector is found to be +2 V. For $V_{CC} = +5$ V and $R_C = 1$ kΩ, find I_C and the small-signal voltage gain. For a change $\Delta v_{BE} = +5$ mV, calculate the resulting Δv_O. Calculate it two ways: by using the transistor exponential characteristic Δi_C, and approximately, using the small-signal voltage gain. Repeat for $\Delta v_{BE} = -5$ mV. Summarize your results in a table.

***7.19** Consider the amplifier circuit of Fig. 7.6 when operated with a supply voltage $V_{CC} = +3$V.

(a) What is the theoretical maximum voltage gain that this amplifier can provide?
(b) What value of V_{CE} must this amplifier be biased at to provide a voltage gain of –60 V/V?
(c) If the dc collector current I_C at the bias point in (b) is to be 0.5 mA, what value of R_C should be used?
(d) What is the value of V_{BE} required to provide the bias point mentioned above? Assume that the BJT has $I_S = 10^{-15}$ A.
(e) If a sine-wave signal v_{be} having a 5-mV peak amplitude is superimposed on V_{BE}, find the corresponding output voltage signal v_{ce} that will be superimposed on V_{CE} assuming linear operation around the bias point.
(f) Characterize the signal current i_c that will be superimposed on the dc bias current I_C.

(g) What is the value of the dc base current I_B at the bias point? Assume $\beta = 100$. Characterize the signal current i_b that will be superimposed on the base current I_B.

(h) Dividing the amplitude of v_{be} by the amplitude of i_b, evaluate the incremental (or small-signal) input resistance of the amplifier.

(i) Sketch and clearly label correlated graphs for v_{BE}, v_{CE}, i_C, and i_B versus time. Note that each graph consists of a dc or average value and a superimposed sine wave. Be careful of the phase relationships of the sine waves.

7.20 The essence of transistor operation is that a change in v_{BE}, Δv_{BE}, produces a change in i_C, Δi_C. By keeping Δv_{BE} small, Δi_C is approximately linearly related to Δv_{BE}, $\Delta i_C = g_m \Delta v_{BE}$, where g_m is known as the transistor transconductance. By passing Δi_C through R_C, an output voltage signal Δv_O is obtained. Use the expression for the small-signal voltage gain in Eq. (7.20) to derive an expression for g_m. Find the value of g_m for a transistor biased at $I_C = 0.5$ mA.

7.21 The purpose of this problem is to illustrate the application of graphical analysis to the circuit shown in Fig. P7.21. Sketch $i_C - v_{CE}$ characteristic curves for the BJT for $i_B = 10\ \mu$A, 20 μA, 30 μA, and 40 μA. Assume the lines to be horizontal (i.e., neglect the Early effect), and let $\beta = 100$. For $V_{CC} = 5$ V and $R_C = 1$ kΩ, sketch the load line. What peak-to-peak collector voltage swing will result for i_B varying over the range 10 μA to 40 μA? If the BJT is biased at $V_{CE} = \frac{1}{2} V_{CC}$, find the value of I_C and I_B. If at this current $V_{BE} = 0.7$ V and if $R_B = 100$ kΩ, find the required value of V_{BB}.

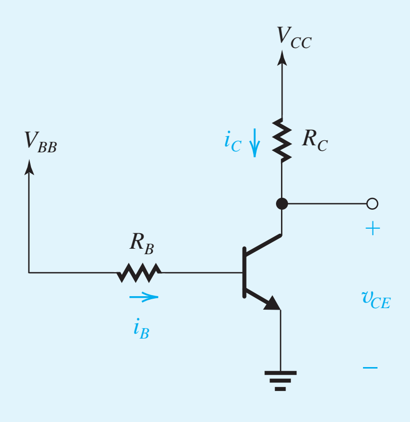


Figure P7.21

***7.22** Sketch the $i_C - v_{CE}$ characteristics of an *npn* transistor having $\beta = 100$ and $V_A = 100$ V. Sketch characteristic curves for $i_B = 20\ \mu$A, 50 μA, 80 μA, and 100 μA. For the purpose of this sketch, assume that $i_C = \beta i_B$ at $v_{CE} = 0$. Also, sketch the load line obtained for $V_{CC} = 10$ V and $R_C = 1$ kΩ. If the dc bias current into the base is 50 μA, write the equation for the corresponding $i_C - v_{CE}$ curve. Also, write the equation for the load line, and solve the two equations to obtain V_{CE} and I_C. If the input signal causes a sinusoidal signal of 30-μA peak amplitude to be superimposed on I_B, find the corresponding signal components of i_C and v_{CE}.

Section 7.2: Small-Signal Operation and Models

***7.23** This problem investigates the nonlinear distortion introduced by a MOSFET amplifier. Let the signal v_{gs} be a sine wave with amplitude V_{gs}, and substitute $v_{gs} = V_{gs} \sin \omega t$ in Eq. (7.28). Using the trigonometric identity $\sin^2\theta = \frac{1}{2} - \frac{1}{2}\cos 2\theta$, show that the ratio of the signal at frequency 2ω to that at frequency ω, expressed as a percentage (known as the second-harmonic distortion) is

$$\text{Second-harmonic distortion} = \frac{1}{4}\frac{V_{gs}}{V_{OV}} \times 100$$

If in a particular application V_{gs} is 10 mV, find the minimum overdrive voltage at which the transistor should be operated so that the second-harmonic distortion is kept to less than 1%.

7.24 Consider an NMOS transistor having $k_n = 10$ mA/V^2. Let the transistor be biased at $V_{OV} = 0.2$ V. For operation in saturation, what dc bias current I_D results? If a 0.02-V signal is superimposed on V_{GS}, find the corresponding increment in collector current by evaluating the total collector current i_D and subtracting the dc bias current I_D. Repeat for a –0.02-V signal. Use these results to estimate g_m of the FET at this bias point. Compare with the value of g_m obtained using Eq. (7.33).

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

7.25 Consider the FET amplifier of Fig. 7.10 for the case $V_t = 0.4$ V, $k_n = 5$ mA/V^2, $V_{GS} = 0.6$ V, $V_{DD} = 1.8$ V, and $R_D = 10$ kΩ.

(a) Find the dc quantities I_D and V_{DS}.
(b) Calculate the value of g_m at the bias point.
(c) Calculate the value of the voltage gain.
(d) If the MOSFET has $\lambda = 0.1$ V^{-1}, find r_o at the bias point and calculate the voltage gain.

D *7.26 An NMOS amplifier is to be designed to provide a 0.20-V peak output signal across a 20-kΩ load that can be used as a drain resistor. If a gain of at least 10 V/V is needed, what g_m is required? Using a dc supply of 1.8 V, what values of I_D and V_{OV} would you choose? What W/L ratio is required if $\mu_n C_{ox} = 200\ \mu$A/V^2? If $V_t = 0.4$ V, find V_{GS}.

D *7.27 In this problem we investigate an optimum design of the CS amplifier circuit of Fig. 7.10. First, use the voltage gain expression $A_v = -g_m R_D$ together with Eq. (7.42) for g_m to show that

$$A_v = -\frac{2I_D R_D}{V_{OV}} = -\frac{2(V_{DD} - V_{DS})}{V_{OV}}$$

Next, let the maximum positive input signal be $\hat{v}_i$. To keep the second-harmonic distortion to an acceptable level, we bias the MOSFET to operate at an overdrive voltage $V_{OV} \gg \hat{v}_i$. Let $V_{OV} = m\hat{v}_i$. Now, to maximize the voltage gain $|A_v|$, we design for the lowest possible V_{DS}. Show that the minimum V_{DS} that is consistent with allowing a negative signal voltage swing at the drain of $|A_v|\ \hat{v}_i$ while maintaining saturation-mode operation is given by

$$V_{DS} = \frac{V_{OV} + \hat{v}_i + 2V_{DD}(\hat{v}_i/V_{OV})}{1 + 2(\hat{v}_i/V_{OV})}$$

Now, find V_{OV}, V_{DS}, A_v, and $\hat{v}_o$ for the case $V_{DD} = 2.5$ V, $\hat{v}_i = 20$ mV, and $m = 15$. If it is desired to operate this transistor at $I_D = 200\ \mu$A, find the values of R_D and W/L, assuming that for this process technology $k_n' = 100\ \mu$A/V^2.

7.28 In the table below, for MOS transistors operating under a variety of conditions, complete as many entries as possible. Although some data is not available, it is always possible to calculate g_m using one of Eqs. (7.40), (7.41), or (7.42). Assume $\mu_n = 500$ cm^2/V·s, $\mu_p = 250$ cm^2/V·s, and $C_{ox} = 0.4$ fF/μm^2.

			Voltages (V)			Dimensions (μm)				
Case	**Type**	**I_D (mA)**	**$\lvert V_{GS}\rvert$**	**$\lvert V_t\rvert$**	**V_{OV}**	**W**	**L**	**W/L**	**$k'(W/L)$**	**g_m(mA/V)**
a	N	1	3	2			1			
b	N	1		0.7	0.5	50				
c	N	10			2		1			
d	N	0.5			0.5					
e	N	0.1				10	2			
f	N		1.8	0.8		40	4			
g	P	0.5						25		
h	P		3	1					0.5	
i	P	10				4000	2			
j	P	10			4					
k	P				1	30	3			
l	P				5				0.08	

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

7.29 An NMOS technology has $\mu_n C_{ox} = 250\ \mu\text{A/V}^2$ and $V_t = 0.5$ V. For a transistor with $L = 0.5\ \mu$m, find the value of W that results in $g_m = 2$ mA/V at $I_D = 0.25$ mA. Also, find the required V_{GS}.

7.30 For the NMOS amplifier in Fig. P7.30, replace the transistor with its T equivalent circuit, assuming $\lambda = 0$. Derive expressions for the voltage gains v_s/v_i and v_d/v_i.

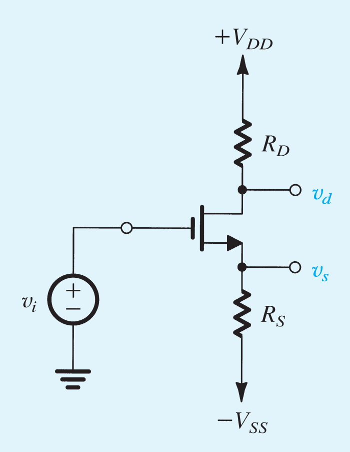


Figure P7.30

SIM **7.31** In the circuit of Fig. P7.31, the NMOS transistor has $|V_t| = 0.5$ V and $V_A = 50$ V and operates with $V_D = 1$ V. What is the voltage gain v_o/v_i ? What do V_D and the gain become for I increased to 1 mA?

7.32 For a 0.18-μm CMOS fabrication process: $V_{tn} = 0.5$ V, $V_{tp} = -0.5$ V, $\mu_n C_{ox} = 400\ \mu\text{A/V}^2$, $\mu_p C_{ox} = 100\ \mu\text{A/V}^2$, $C_{ox} = 8.6\ \text{fF}/\mu\text{m}^2$, V_A (n-channel devices) $= 5L$ (μm), and $|V_A|$ (p-channel devices) $= 6L$ (μm). Find the small-signal model parameters (g_m and r_o) for both an NMOS and a PMOS transistor having $W/L = 10\ \mu\text{m}/0.5\ \mu\text{m}$ and operating at $I_D = 100\ \mu$A. Also, find the overdrive voltage at which each device must be operating.

***7.33** Figure P7.33 shows a discrete-circuit amplifier. The input signal v_{sig} is coupled to the gate through a very large capacitor (shown as infinite). The transistor source is connected to ground at signal frequencies via a very large capacitor (shown as infinite). The output voltage signal that develops at the drain is coupled to a load resistance via a very large capacitor (shown as infinite). All capacitors behave as short circuits for signals and as open circuits for dc.

(a) If the transistor has $V_t = 1$ V, and $k_n = 4\ \text{mA/V}^2$, verify that the bias circuit establishes $V_{GS} = 1.5$ V, $I_D = 0.5$ mA, and $V_D = +7.0$ V. That is, assume these values, and verify that they are consistent with the values of the circuit components and the device parameters.
(b) Find g_m and r_o if $V_A = 100$ V.
(c) Draw a complete small-signal equivalent circuit for the amplifier, assuming all capacitors behave as short circuits at signal frequencies.
(d) Find R_{in}, v_{gs}/v_{sig}, v_o/v_{gs}, and v_o/v_{sig}.

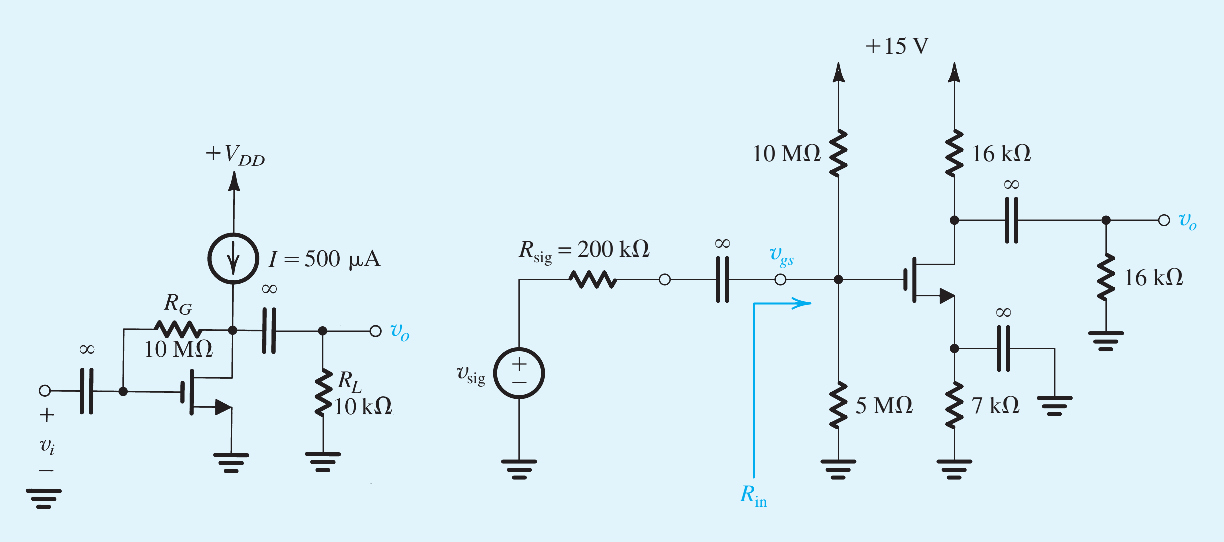


Figure P7.31

Figure P7.33

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

7.34 Consider a transistor biased to operate in the active mode at a dc collector current I_C. Calculate the collector signal current as a fraction of I_C (i.e., i_c/I_C) for input signals v_{be} of +1 mV, –1 mV, +2 mV, –2 mV, +5 mV, –5 mV, +8 mV, –8 mV, +10 mV, –10 mV, +12 mV, and –12 mV. In each case do the calculation two ways:

(a) using the exponential characteristic, and
(b) using the small-signal approximation.

Present your results in the form of a table that includes a column for the error introduced by the small-signal approximation. Comment on the range of validity of the small-signal approximation.

7.35 An *npn* BJT with grounded emitter is operated with $V_{BE} = 0.700$ V, at which the collector current is 0.5 mA. A 5-kΩ resistor connects the collector to a +5-V supply. What is the resulting collector voltage V_C? Now, if a signal applied to the base raises v_{BE} to 705 mV, find the resulting total collector current i_C and total collector voltage v_C using the exponential i_C–v_{BE} relationship. For this situation, what are v_{be} and v_c? Calculate the voltage gain v_c/v_{be}. Compare with the value obtained using the small-signal approximation, that is, $-g_m R_C$.

7.36 A transistor with $\beta = 100$ is biased to operate at a dc collector current of 0.5 mA. Find the values of g_m, r_π, and r_e. Repeat for a bias current of 50 µA.

7.37 A *pnp* BJT is biased to operate at $I_C = 1.0$ mA. What is the associated value of g_m? If $\beta = 100$, what is the value of the small-signal resistance seen looking into the emitter (r_e)? Into the base (r_π)? If the collector is connected to a 5-kΩ load, with a signal of 5-mV peak applied between base and emitter, what output signal voltage results?

D 7.38 A designer wishes to create a BJT amplifier with a g_m of 30 mA/V and a base input resistance of 3000 Ω or more. What collector-bias current should he choose? What is the minimum β he can tolerate for the transistor used?

7.39 A transistor operating with nominal g_m of 40 mA/V has a β that ranges from 50 to 150. Also, the bias circuit, being less than ideal, allows a ±20% variation in I_C. What are the extreme values found of the resistance looking into the base?

7.40 In the circuit of Fig. 7.20, V_{BE} is adjusted so that $V_C = 1$ V. If $V_{CC} = 3$ V, $R_C = 2$ kΩ, and a signal $v_{be} = 0.005 \sin \omega t$ volts is applied, find expressions for the total instantaneous quantities $i_C(t)$, $v_C(t)$, and $i_B(t)$. The transistor has $\beta = 100$. What is the voltage gain?

D *7.41 We wish to design the amplifier circuit of Fig. 7.20 under the constraint that V_{CC} is fixed. Let the input signal $v_{be} = \hat{V}_{be} \sin \omega t$, where $\hat{V}_{be}$ is the maximum value for acceptable linearity. For the design that results in the largest signal at the collector, without the BJT leaving the active region, show that

$$R_C I_C = (V_{CC} - 0.3) \Big/ \left(1 + \frac{\hat{V}_{be}}{V_T}\right)$$

and find an expression for the voltage gain obtained. For $V_{CC} = 3$ V and $\hat{V}_{be} = 5$ mV, find the dc voltage at the collector, the amplitude of the output voltage signal, and the voltage gain.

7.42 The table below summarizes some of the basic attributes of a number of BJTs of different types, operating as amplifiers under various conditions. Provide the missing entries. (*Note:* Isn't it remarkable how much two parameters can reveal?)

7.43 A BJT is biased to operate in the active mode at a dc collector current of 1 mA. It has a β of 100 and V_A of 100 V. Give the four small-signal models (Figs. 7.25 and 7.27) of the BJT complete with the values of their parameters.

Transistor	a	b	c	d	e	f	g
α	1.000					0.90	
β		100		∞			
I_C (mA)	1.00		1.00				
I_E (mA)		1.00				5	
I_B (mA)			0.020				1.10
g_m (mA/V)							700
r_e (Ω)				25	100		
r_π (Ω)					10.1 kΩ		

7.44 Using the T model of Fig. 7.26(a), show that the input resistance between base and emitter, looking into the base, is equal to r_π.

7.45 Show that the collector current provided by the model of Fig. 7.26(b) is equal to that provided by the model in Fig. 7.26(a).

7.46 Show that the hybrid-π model of Fig. 7.24(b) is the incremental version of the large-signal model of Fig. 6.5(d).

7.47 Show that the T model of Fig. 7.26(b) is the incremental version of the large-signal model of Fig. 6.5(b).

7.48 The transistor amplifier in Fig. P7.48 is biased with a current source I and has a very high β. Find the dc voltage at the collector, V_C. Also, find the value of r_e. Replace the transistor with the T model of Fig. 7.26(b) (note that the dc current source I should be replaced with an open circuit). Hence find the voltage gain v_c/v_i.

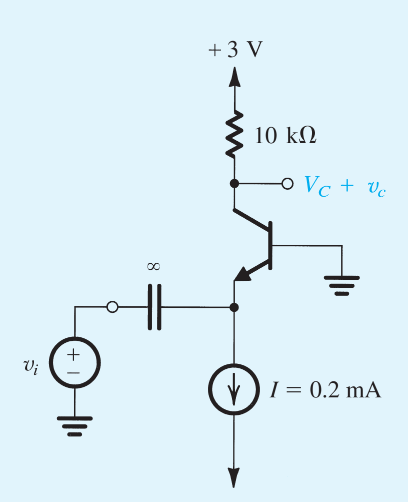


Figure P7.48

7.49 For the conceptual circuit shown in Fig. 7.23, $R_C = 2\ \text{k}\Omega$, $g_m = 50$ mA/V, and $\beta = 100$. If a peak-to-peak output voltage of 1 V is measured at the collector, what are the peak-to-peak values of v_{be} and i_b?

7.50 Figure P7.50 shows the circuit of an amplifier fed with a signal source v_{sig} with a source resistance R_{sig}. The bias circuitry is not shown. Replace the BJT with its hybrid-π equivalent circuit of Fig. 7.24(a). Find the input resistance $R_{in} \equiv v_\pi/i_b$, the voltage transmission from source to amplifier input, v_π/v_{sig}, and the voltage gain from base to collector, v_o/v_π. Use these to show that the overall voltage gain v_o/v_{sig} is given by

$$\frac{v_o}{v_{sig}} = -\frac{\beta R_C}{r_\pi + R_{sig}}$$

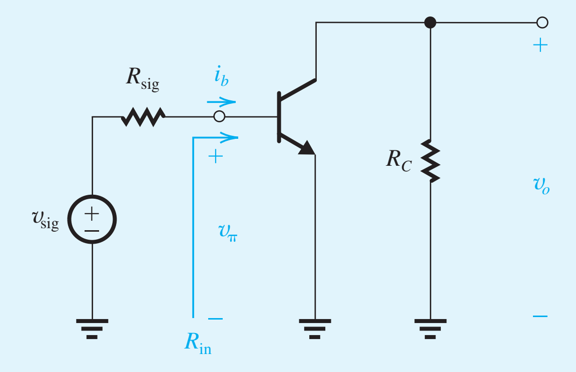


Figure P7.50

7.51 Figure P7.51 shows a transistor with the collector connected to the base. The bias arrangement is not shown. Since a zero v_{BC} implies operation in the active mode, the BJT can be replaced by one of the small-signal models of Figs. 7.24 and 7.26. Use the model of Fig. 7.26(b) and show that the resulting two-terminal device, known as a diode-connected transistor, has a small-signal resistance r equal to r_e.

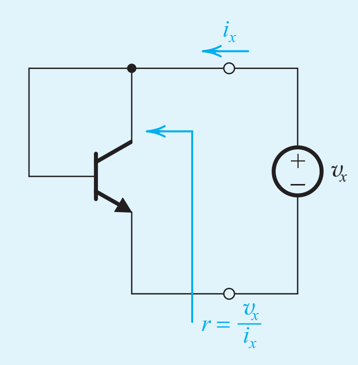


Figure P7.51

7.52 Figure P7.52 shows a particular configuration of BJT amplifiers known as "emitter follower." The bias arrangement is not shown. Replace the BJT with its T equivalent-circuit

model of Fig. 7.26(b). Show that

$$R_{in} \equiv \frac{v_i}{i_b} = (\beta + 1)(r_e + R_e)$$

$$\frac{v_o}{v_i} = \frac{R_e}{R_e + r_e}$$

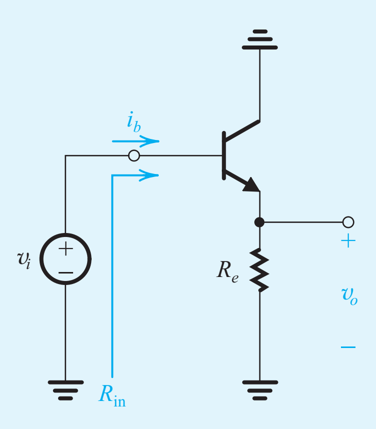


Figure P7.52

7.53 For the circuit shown in Fig. P7.53, draw a complete small-signal equivalent circuit utilizing an appropriate T model for the BJT (use $\alpha = 0.99$). Your circuit should show the values of all components, including the model parameters. What is the input resistance R_{in}? Calculate the overall voltage gain (v_o/v_{sig}).

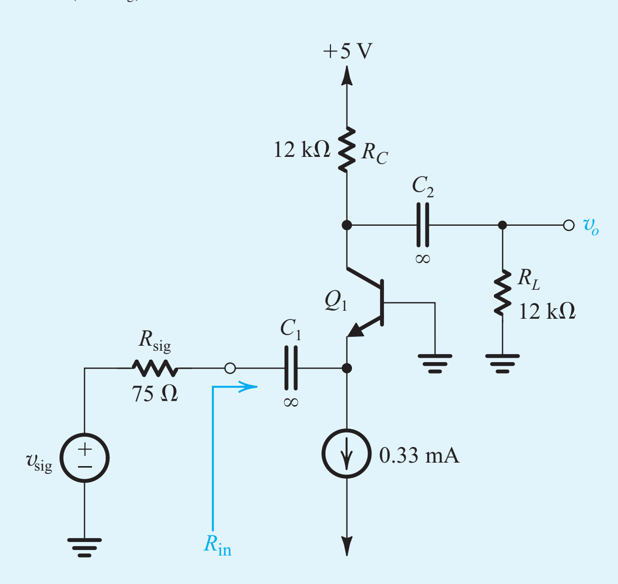


Figure P7.53

7.54 In the circuit shown in Fig. P7.54, the transistor has a β of 200. What is the dc voltage at the collector? Replacing the BJT with one of the hybrid-π models (neglecting r_o), draw the equivalent circuit of the amplifier. Find the input resistances R_{ib} and R_{in} and the overall voltage gain (v_o/v_{sig}). For an output signal of ±0.4 V, what values of v_{sig} and v_b are required?

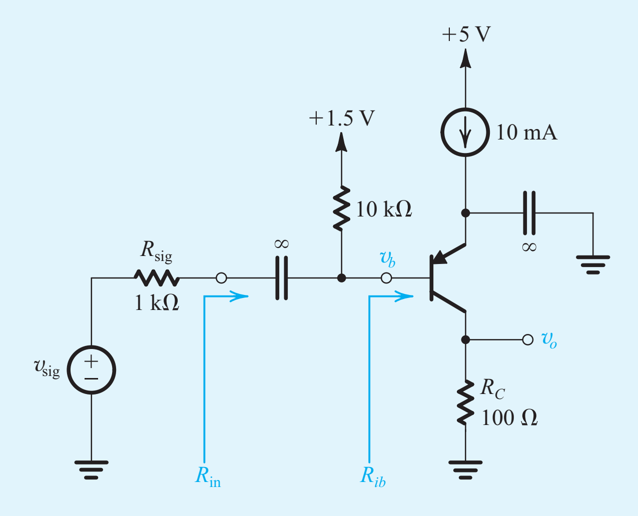


Figure P7.54

7.55 Consider the augmented hybrid-π model shown in Fig. 7.25(a). Disregarding how biasing is to be done, what is the largest possible voltage gain available for a signal source connected directly to the base and a very-high-resistance load? Calculate the value of the maximum possible gain for $V_A = 25$ V and $V_A = 125$ V.

D 7.56 Redesign the circuit of Fig. 7.30(a) by raising the resistor values by a factor n to increase the resistance seen by the input v_i to 75 Ω. What value of voltage gain results? Grounded-base circuits of this kind are used in systems such as cable TV, in which, for highest-quality signaling, load resistances need to be "matched" to the equivalent resistances of the interconnecting cables.

D *7.57 Design an amplifier using the configuration of Fig. 7.30(a). The power supplies available are ±5 V. The input signal source has a resistance of 50 Ω, and it is required that the amplifier input resistance match this value. (Note that $R_{in} = r_e \| R_E \simeq r_e$.) The amplifier is to have the greatest possible voltage gain and the largest possible output signal but retain small-signal linear operation (i.e., the signal component across the base–emitter junction should be limited to no more

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

than 10 mV). Find appropriate values for R_E and R_C. What is the value of voltage gain realized from signal source to output?

***7.58** The transistor in the circuit shown in Fig. P7.58 is biased to operate in the active mode. Assuming that β is very large, find the collector bias current I_C. Replace the transistor with the small-signal equivalent-circuit model of Fig. 7.26(b) (remember to replace the dc power supply with a short circuit). Analyze the resulting amplifier equivalent circuit to show that

$$\frac{v_{o1}}{v_i} = \frac{R_E}{R_E + r_e}$$

$$\frac{v_{o2}}{v_i} = \frac{-\alpha R_C}{R_E + r_e}$$

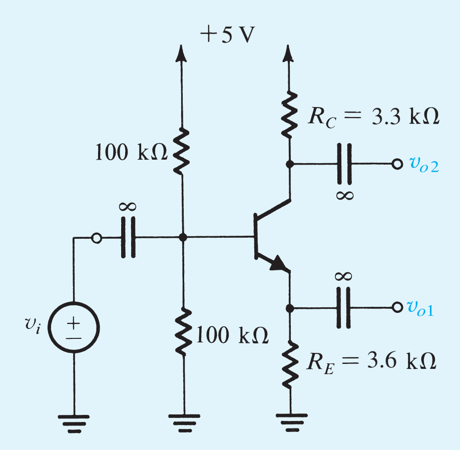


Figure P7.58

Find the values of these voltage gains (for $\alpha \simeq 1$). Now, if the terminal labeled v_{o1} is connected to ground, what does the voltage gain v_{o2}/v_i become?

Section 7.3: Basic Configurations

7.59 An amplifier with an input resistance of 100 kΩ, an open-circuit voltage gain of 100 V/V, and an output resistance of 100 Ω is connected between a 20-kΩ signal source and a 2-kΩ load. Find the overall voltage gain G_v. Also find the current gain, defined as the ratio of the load current to the current drawn from the signal source.

D 7.60 Specify the parameters R_{in}, A_{vo}, and R_o of an amplifier that is to be connected between a 100-kΩ source and a 2-kΩ load and is required to meet the following specifications:

(a) No more than 5% of the signal strength is lost in the connection to the amplifier input;
(b) If the load resistance changes from the nominal value of 2 kΩ to a low value of 1 kΩ, the change in output voltage is limited to 5% of nominal value; and
(c) The nominal overall voltage gain is 10 V/V.

7.61 Figure P7.61 shows an alternative equivalent-circuit representation of an amplifier. If this circuit is to be equivalent to that in Fig. 7.34(b) show that $G_m = A_{vo}/R_o$. Also convince yourself that the transconductance G_m is defined as

$$G_m = \left.\frac{i_o}{v_i}\right|_{R_L=0}$$

and hence is known as the short-circuit transconductance. Now, if the amplifier is fed with a signal source (v_{sig}, R_{sig}) and is connected to a load resistance R_L show that the gain of the amplifier proper A_v is given by $A_v = G_m(R_o \parallel R_L)$ and the overall voltage gain G_v is given by

$$G_v = \frac{R_{in}}{R_{in} + R_{sig}} G_m(R_o \parallel R_L)$$

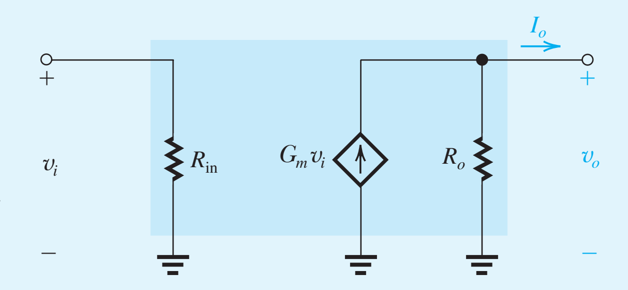


Figure P7.61

7.62 An alternative equivalent circuit of an amplifier fed with a signal source (v_{sig}, R_{sig}) and connected to a load R_L is shown in Fig. P7.62. Here G_{vo} is the open-circuit overall voltage gain,

$$G_{vo} = \left.\frac{v_o}{v_{sig}}\right|_{R_L=\infty}$$

and R_{out} is the output resistance with v_{sig} set to zero. This is different than R_o. Show that

$$G_{vo} = \frac{R_i}{R_i + R_{sig}} A_{vo}$$

where $R_i = R_{in}|_{R_L=\infty}$.

Also show that the overall voltage gain is

$$G_v = G_{vo} \frac{R_L}{R_L + R_{out}}$$

****7.63** Most practical amplifiers have internal feedback that make them non-unilateral. In such a case, R_{in} depends on R_L. To illustrate this point we show in Fig. P7.63 the equivalent circuit of an amplifier where a feedback resistance R_f models the internal feedback mechanism that is present in this amplifier. It is R_f that makes the amplifier non-unilateral. Show that

$$R_{in} = R_1 \| \left[\frac{R_f + (R_2 \| R_L)}{1 + g_m(R_2 \| R_L)} \right]$$

$$A_{vo} = -g_m R_2 \frac{1 - 1/(g_m R_f)}{1 + (R_2/R_f)}$$

$$R_o = R_2 \| R_f$$

Evaluate R_{in}, A_{vo}, and R_o for the case $R_1 = 100\text{ k}\Omega$, $R_f = 1\text{ M}\Omega$, $g_m = 100\text{ mA/V}$, $R_2 = 100\ \Omega$, and $R_L = 1\text{ k}\Omega$. Which of the amplifier characteristic parameters is most affected by R_f (that is, relative to the case with $R_f = \infty$)? For $R_{sig} = 100\text{ k}\Omega$ determine the overall voltage gain, G_v, with and without R_f present.

7.64 Calculate the overall voltage gain of a CS amplifier fed with a 1-MΩ source and connected to a 10-kΩ load. The MOSFET has $g_m = 2\text{ mA/V}$, and a drain resistance $R_D = 10\text{ k}\Omega$ is utilized.

7.65 A CS amplifier utilizes a MOSFET with $\mu_n C_{ox} = 400\ \mu\text{A/V}^2$ and $W/L = 10$. It is biased at $I_D = 320\ \mu\text{A}$ and uses $R_D = 10\text{ k}\Omega$. Find R_{in}, A_{vo}, and R_o. Also, if a load resistance of 10 kΩ is connected to the output, what overall voltage gain G_v is realized? Now, if a 0.2-V peak sine-wave signal is required at the output, what must the peak amplitude of v_{sig} be?

7.66 A common-source amplifier utilizes a MOSFET operated at $V_{OV} = 0.25$ V. The amplifier feeds a load resistance $R_L = 15\text{ k}\Omega$. The designer selects $R_D = 2R_L$. If it is required to realize an overall voltage gain G_v of -10 V/V what g_m is needed? Also specify the bias current I_D. If, to increase the output signal swing, R_D is reduced to $R_D = R_L$, what does G_v become?

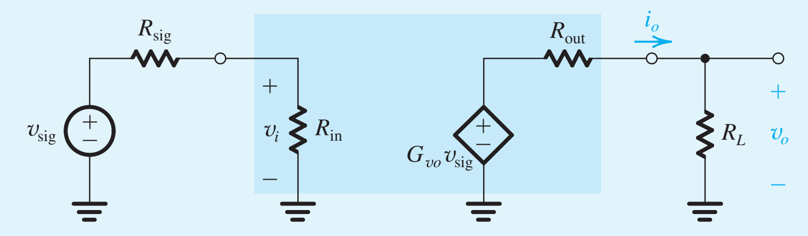


Figure P7.62

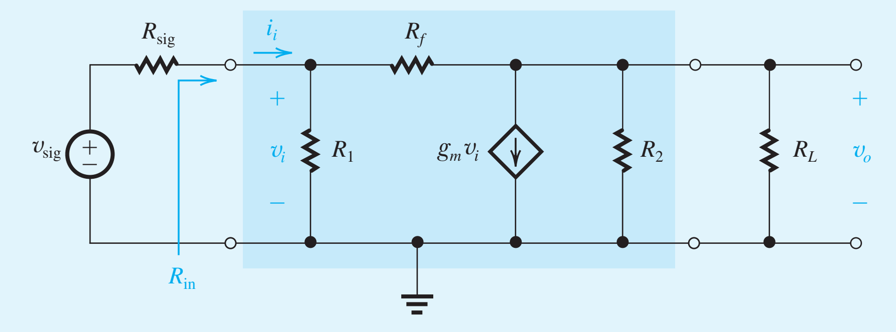


Figure P7.63

7.67 Two identical CS amplifiers are connected in cascade. The first stage is fed with a source v_{sig} having a resistance $R_{sig} = 200\ \text{k}\Omega$. A load resistance $R_L = 10\ \text{k}\Omega$ is connected to the drain of the second stage. Each MOSFET is biased at $I_D = 0.3$ mA and operates with $V_{OV} = 0.2$ V. Each stage utilizes a drain resistance $R_D = 10\ \text{k}\Omega$.

(a) Sketch the equivalent circuit of the two-stage amplifier.
(b) Calculate the overall voltage gain G_v.

7.68 A CE amplifier utilizes a BJT with $\beta = 100$ biased at $I_C = 0.5$ mA; it has a collector resistance $R_C = 10\ \text{k}\Omega$. Find R_{in}, R_o, and A_{vo}. If the amplifier is fed with a signal source having a resistance of $10\ \text{k}\Omega$, and a load resistance $R_L = 10\ \text{k}\Omega$ is connected to the output terminal, find the resulting A_v and G_v. If the peak voltage of the sine wave appearing between base and emitter is to be limited to 5 mV, what $\hat{v}_{sig}$ is allowed, and what output voltage signal appears across the load?

D *7.69 In this problem we investigate the effect of the inevitable variability of β on the realized gain of the CE amplifier. For this purpose, use the overall gain expression in Eq. (7.114).

$$|G_v| = \frac{R_L'}{(R_{sig}/\beta) + (1/g_m)}$$

where $R_L' = R_L \parallel R_C$.

Consider the case $R_L' = 10\ \text{k}\Omega$ and $R_{sig} = 10\ \text{k}\Omega$, and let the BJT be biased at $I_C = 1$ mA. The BJT has a nominal β of 100.

(a) What is the nominal value of $|G_v|$?
(b) If β can be anywhere between 50 and 150, what is the corresponding range of $|G_v|$?
(c) If in a particular design, it is required to maintain $|G_v|$ within $\pm 20\%$ of its nominal value, what is the maximum allowable range of β?
(d) If it is not possible to restrict β to the range found in (c), and the designer has to contend with β in the range 50 to 150, what value of bias current I_C would result in $|G_v|$ falling in a range of $\pm 20\%$ of a new nominal value? What is the nominal value of $|G_v|$ in this case?

7.70 Two identical CE amplifiers are connected in cascade. The first stage is fed with a source v_{sig} having a resistance $R_{sig} = 10\ \text{k}\Omega$. A load resistance $R_L = 10\ \text{k}\Omega$ is connected to the collector of the second stage. Each BJT is biased at $I_C = 0.25$ mA and has $\beta = 100$. Each stage utilizes a collector resistance $R_C = 10\ \text{k}\Omega$.

(a) Sketch the equivalent circuit of the two-stage amplifier.
(b) Find the overall voltage gain, v_{o2}/v_{sig}.

7.71 A MOSFET connected in the CS configuration has a transconductance $g_m = 5$ mA/V. When a resistance R_s is connected in the source lead, the effective transconductance is reduced to 2 mA/V. What do you estimate the value of R_s to be?

7.72 A CS amplifier using an NMOS transistor with $g_m = 2$ mA/V is found to have an overall voltage gain of -10 V/V. What value should a resistance R_s inserted in the source lead have to reduce the overall voltage gain to -5 V/V?

7.73 The overall voltage gain of a CS amplifier with a resistance $R_s = 0.5\ \text{k}\Omega$ in the source lead was measured and found to be -10 V/V. When R_s was shorted, but the circuit operation remained linear, the gain doubled. What must g_m be? What value of R_s is needed to obtain an overall voltage gain of -16 V/V?

7.74 A CE amplifier utilizes a BJT with $\beta = 100$ biased at $I_C = 0.5$ mA and has a collector resistance $R_C = 12\ \text{k}\Omega$ and a resistance $R_e = 250\ \Omega$ connected in the emitter. Find R_{in}, A_{vo}, and R_o. If the amplifier is fed with a signal source having a resistance of $10\ \text{k}\Omega$, and a load resistance $R_L = 12\ \text{k}\Omega$ is connected to the output terminal, find the resulting A_v and G_v. If the peak voltage of the sine wave appearing between base and emitter is to be limited to 5 mV, what $\hat{v}_{sig}$ is allowed, and what output voltage signal appears across the load?

D 7.75 Design a CE amplifier with a resistance R_e in the emitter to meet the following specifications:

(i) Input resistance $R_{in} = 15\ \text{k}\Omega$.
(ii) When fed from a signal source with a peak amplitude of 0.15 V and a source resistance of $30\ \text{k}\Omega$, the peak amplitude of v_π is 5 mV.

Specify R_e and the bias current I_C. The BJT has $\beta = 74$. If the total resistance in the collector is $6\ \text{k}\Omega$, find the overall voltage gain G_v and the peak amplitude of the output signal v_o.

SIM **D 7.76** Inclusion of an emitter resistance R_e reduces the variability of the gain G_v due to the inevitable wide variance in the value of β. Consider a CE amplifier operating between a signal source with $R_{sig} = 10\ \text{k}\Omega$ and a total collector resistance $R_C \parallel R_L$ of $10\ \text{k}\Omega$. The BJT is biased at $I_C = 1$ mA and its β is specified to be nominally 100 but can lie in the range of 50 to 150. First determine the nominal value and the range of

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

$|G_v|$ without resistance R_e. Then select a value for R_e that will ensure that $|G_v|$ be within ±20% of its new nominal value. Specify the value of R_e, the new nominal value of $|G_v|$, and the expected range of $|G_v|$.

7.77 A CG amplifier using an NMOS transistor for which $g_m = 2$ mA/V has a 5-kΩ drain resistance R_D and a 5-kΩ load resistance R_L. The amplifier is driven by a voltage source having a 750-Ω resistance. What is the input resistance of the amplifier? What is the overall voltage gain G_v ? By what factor must the bias current I_D of the MOSFET be changed so that R_{in} matches R_{sig}?

7.78 A CG amplifier when fed with a signal source having $R_{sig} = 100\ \Omega$ is found to have an overall voltage gain of 12 V/V. When a 100-Ω resistance was added in series with the signal generator the overall voltage gain decreased to 10 V/V. What must g_m of the MOSFET be? If the MOSFET is biased at $I_D = 0.25$ mA, at what overdrive voltage must it be operating?

D 7.79 A CB amplifier is operating with $R_L = 10$ kΩ, $R_C = 10$ kΩ, and $R_{sig} = 50\ \Omega$. At what current I_C should the transistor be biased for the input resistance R_{in} to equal that of the signal source? What is the resulting overall voltage gain? Assume $\alpha \simeq 1$.

7.80 For the circuit in Fig. P7.80, let $R_{sig} \gg r_e$ and $\alpha \simeq 1$. Find v_o.

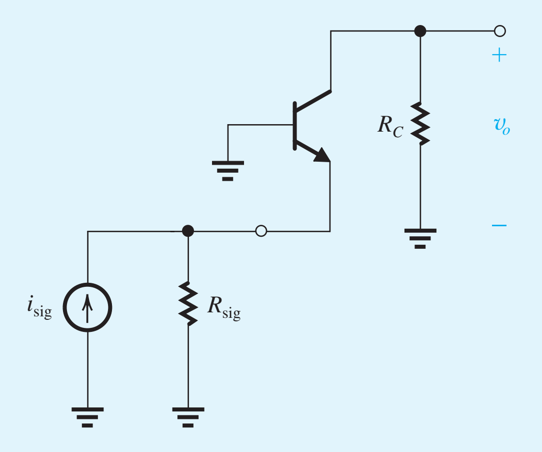


Figure P7.80

7.81 A CB amplifier is biased at $I_E = 0.2$ mA with $R_C = R_L = 10$ kΩ and is driven by a signal source with $R_{sig} = 0.5$ kΩ. Find the overall voltage gain G_v. If the maximum signal amplitude of the voltage between base and emitter is limited to 10 mV, what are the corresponding amplitudes of v_{sig} and v_o? Assume $\alpha \simeq 1$.

7.82 A source follower is required to connect a high-resistance source to a load whose resistance is nominally 2 kΩ but can be as low as 1.5 kΩ and as high as 5 kΩ. What is the maximum output resistance that the source follower must have if the output voltage is to remain within ±10% of nominal value? If the MOSFET has $k_n = 2.5$ mA/V^2, at what current I_D must it be biased? At what overdrive voltage is the MOSFET operating?

D 7.83 A source follower is required to deliver a 0.5-V peak sinusoid to a 2-kΩ load. If the peak amplitude of v_{gs} is to be limited to 50 mV, and the MOSFET transconductance parameter k_n is 5 mA/V^2, what is the lowest value of I_D at which the MOSFET can be biased? At this bias current, what are the maximum and minimum currents that the MOSFET will be conducting (at the positive and negative peaks of the output sine wave)? What must the peak amplitude of v_{sig} be?

D 7.84 An emitter follower is required to deliver a 0.5-V peak sinusoid to a 2-kΩ load. If the peak amplitude of v_{be} is to be limited to 5 mV, what is the lowest value of I_E at which the BJT can be biased? At this bias current, what are the maximum and minimum currents that the BJT will be conducting (at the positive and negative peaks of the output sine wave)? If the resistance of the signal source is 200 kΩ, what value of G_v is obtained? Thus determine the required amplitude of v_{sig}. Assume $\beta = 100$.

7.85 An emitter follower with a BJT biased at $I_C = 2$ mA and having $\beta = 100$ is connected between a source with $R_{sig} = 10$ kΩ and a load $R_L = 0.5$ kΩ.

(a) Find R_{in}, v_b/v_{sig}, and v_o/v_{sig}.
(b) If the signal amplitude across the base–emitter junction is to be limited to 10 mV, what is the corresponding amplitude of v_{sig} and v_o?
(c) Find the open-circuit voltage gain G_{vo} and the output resistance R_{out}. Use these values first to verify the value of G_v obtained in (a), then to find the value of G_v obtained with R_L reduced to 250 Ω.

7.86 An emitter follower is operating at a collector bias current of 0.5 mA and is used to connect a 10-kΩ source to a 1-kΩ load. If the nominal value of β is 100, what output resistance R_{out} and overall voltage gain G_v result? Now if

transistor β is specified to lie in the range 50 to 150, find the corresponding range of R_{out} and G_v.

7.87 An emitter follower, when driven from a 5-kΩ source, was found to have an output resistance R_{out} of 150 Ω. The output resistance increased to 250 Ω when the source resistance was increased to 10 kΩ. Find the overall voltage gain when the follower is driven by a 10-kΩ source and loaded by a 1-kΩ resistor.

7.88 For the general amplifier circuit shown in Fig. P7.88 neglect the Early effect.

(a) Find expressions for v_c/v_{sig} and v_e/v_{sig}.
(b) If v_{sig} is disconnected from node X, node X is grounded, and node Y is disconnected from ground and connected to v_{sig}, find the new expression for v_c/v_{sig}.

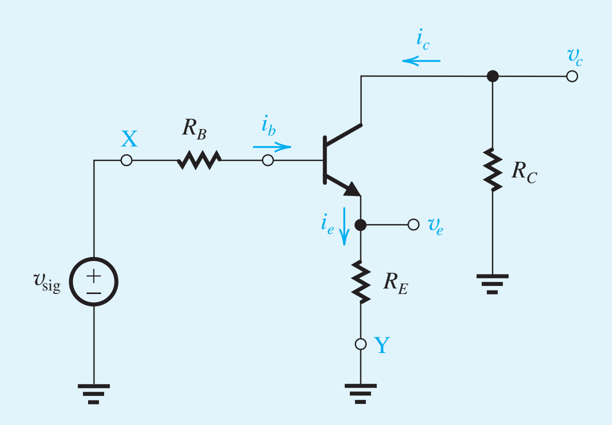


Figure P7.88

7.89 When the Early effect is neglected, the overall voltage gain of a CE amplifier with a collector resistance $R_C = 10$ kΩ is calculated to be −100 V/V. If the BJT is biased at $I_C = 1$ mA and the Early voltage is 100 V, provide a better estimate of the voltage gain G_v.

***7.90** Show that when r_o is taken into account, the voltage gain of the source follower becomes

$$G_v \equiv \frac{v_o}{v_{sig}} = \frac{R_L \| r_o}{(R_L \| r_o) + \dfrac{1}{g_m}}$$

Now, with R_L removed, the voltage gain is carefully measured and found to be 0.98. Then, when R_L is connected and its value is varied, it is found that the gain is halved at $R_L = 500$ Ω. If the amplifier remained linear throughout this measurement, what must the values of g_m and r_o be?

D 7.91 In this problem, we investigate the effect of changing the bias current I_C on the overall voltage gain G_v of a CE amplifier. Consider the situation of a CE amplifier operating with a signal source having $R_{sig} = 10$ kΩ and having $R_C \| R_L = 10$ kΩ. The BJT is specified to have $\beta = 100$ and $V_A = 25$ V. Use Eq. (7.114) (with r_o included in parallel with R_C and R_L in the numerator) to find $|G_v|$ at $I_C = 0.1$ mA, 0.2 mA, 0.5 mA, 1.0 mA, and 1.25 mA. Observe the effect of r_o on limiting $|G_v|$ as I_C is increased. Find the value of I_C that results in $|G_v| = 50$ V/V.

Section 7.4: Biasing

D 7.92 Consider the classical biasing scheme shown in Fig. 7.48(c), using a 9-V supply. For the MOSFET, $V_t = 1$ V, $\lambda = 0$, and $k_n = 2$ mA/V^2. Arrange that the drain current is 1 mA, with about one-third of the supply voltage across each of R_S and R_D. Use 22 MΩ for the larger of R_{G1} and R_{G2}. What are the values of R_{G1}, R_{G2}, R_S, and R_D that you have chosen? Specify them to two significant digits. For your design, how far is the drain voltage from the edge of saturation?

D 7.93 Using the circuit topology displayed in Fig. 7.48(e), arrange to bias the NMOS transistor at $I_D = 0.5$ mA with V_D midway between cutoff and the beginning of triode operation. The available supplies are ±5 V. For the NMOS transistor, $V_t = 1.0$ V, $\lambda = 0$, and $k_n = 1$ mA/V^2. Use a gate-bias resistor of 10 MΩ. Specify R_S and R_D to two significant digits.

D *7.94 In an electronic instrument using the biasing scheme shown in Fig. 7.48(c), a manufacturing error reduces R_S to zero. Let $V_{DD} = 15$ V, $R_{G1} = 10$ MΩ, and $R_{G2} = 5.1$ MΩ. What is the value of V_G created? If supplier specifications allow k_n to vary from 0.2 to 0.3 mA/V^2 and V_t to vary from 1.0 V to 1.5 V, what are the extreme values of I_D that may result? What value of R_S should have been installed to limit the maximum value of I_D to 1.5 mA? Choose an appropriate standard 5% resistor value (refer to Appendix J). What extreme values of current now result?

7.95 An NMOS transistor is connected in the bias circuit of Fig. 7.48(c), with $V_G = 5$ V and $R_S = 3$ kΩ. The transistor has $V_t = 1$ V and $k_n = 2$ mA/V^2. What bias current results? If a transistor for which k_n is 50% higher is used, what is the resulting percentage increase in I_D?

SIM **7.96** The bias circuit of Fig. 7.48(c) is used in a design with $V_G = 5$ V and $R_S = 2$ kΩ. For a MOSFET with

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

$k_n = 2$ mA/V^2, the source voltage was measured and found to be 2 V. What must V_t be for this device? If a device for which V_t is 0.5 V less is used, what does V_S become? What bias current results?

D 7.97 Design the circuit of Fig. 7.48(e) for a MOSFET having $V_t = 1$ V and $k_n = 4$ mA/V^2. Let $V_{DD} = V_{SS} = 5$ V. Design for a dc bias current of 0.5 mA and for the largest possible voltage gain (and thus the largest possible R_D) consistent with allowing a 2-V peak-to-peak voltage swing at the drain. Assume that the signal voltage on the source terminal of the FET is zero.

SIM D 7.98 Design the circuit in Fig. P7.98 so that the transistor operates in saturation with V_D biased 1 V from the edge of the triode region, with $I_D = 1$ mA and $V_D = 3$ V, for each of the following two devices (use a 10-µA current in the voltage divider):

(a) $|V_t| = 1$ V and $k_p' W/L = 0.5$ mA/V^2
(b) $|V_t| = 2$ V and $k_p' W/L = 1.25$ mA/V^2

For each case, specify the values of V_G, V_D, V_S, R_1, R_2, R_S, and R_D.

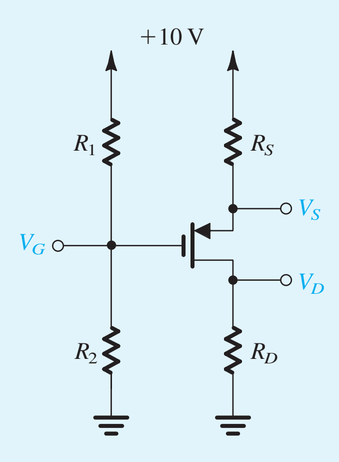


Figure P7.98

D **7.99 A very useful way to characterize the stability of the bias current I_D is to evaluate the sensitivity of I_D relative to a particular transistor parameter whose variability might be large. The sensitivity of I_D relative to the MOSFET parameter $K \equiv \frac{1}{2}k'(W/L)$ is defined as

$$S_K^{I_D} \equiv \frac{\partial I_D / I_D}{\partial K / K} = \frac{\partial I_D}{\partial K}\frac{K}{I_D}$$

and its value, when multiplied by the variability (or tolerance) of K, provides the corresponding expected variability of I_D,

$$\frac{\Delta I_D}{I_D} = S_K^{I_D}\left(\frac{\Delta K}{K}\right)$$

The purpose of this problem is to investigate the use of the sensitivity function in the design of the bias circuit of Fig. 7.48(e).

(a) Show that for V_t constant,

$$S_K^{I_D} = 1 \Big/ \left(1 + 2\sqrt{K I_D} R_S\right)$$

(b) For a MOSFET having $K = 100$ µA/V^2 with a variability of ±10% and $V_t = 1$ V, find the value of R_S that would result in $I_D = 100$ µA with a variability of ±1%. Also, find V_{GS} and the required value of V_{SS}.
(c) If the available supply $V_{SS} = 5$ V, find the value of R_S for $I_D = 100$ µA. Evaluate the sensitivity function, and give the expected variability of I_D in this case.

D **7.100 The variability ($\Delta I_D / I_D$) in the bias current I_D due to the variability ($\Delta V_t / V_t$) in the threshold voltage V_t can be evaluated from

$$\frac{\Delta I_D}{I_D} = S_{V_t}^{I_D}\left(\frac{\Delta V_t}{V_t}\right)$$

where $S_{V_t}^{I_D}$, the sensitivity of I_D relative to V_t, is defined as

$$S_{V_t}^{I_D} = \frac{\partial I_D}{\partial V_t}\frac{V_t}{I_D}$$

(a) For the case of a MOSFET biased with a fixed V_{GS}, show that

$$S_{V_t}^{I_D} = -\frac{2V_t}{V_{OV}}$$

and find the variability in I_D for $V_t = 0.5$ V and $\Delta V_t / V_t = \pm 5\%$. Let the MOSFET be biased at $V_{OV} = 0.25$ V.
(b) For the case of a MOSFET biased with a fixed gate voltage V_G and a resistance R_S included in the source lead, show that

$$S_{V_t}^{I_D} = -\frac{2V_t}{V_{OV} + 2I_D R_S}$$

For the same parameters given in (a), find the required value of ($I_D R_S$) and V_G to limit $\Delta I_D / I_D$ to ±5%. What value of R_S is needed if I_D is 100 µA?

SIM **7.101** In the circuit of Fig. 7.50, let $R_G = 10$ MΩ, $R_D = 10$ kΩ, and $V_{DD} = 10$ V. For each of the following two transistors, find the voltages V_D and V_G.

(a) $V_t = 1$ V and $k_n = 0.5$ mA/V^2
(b) $V_t = 2$ V and $k_n = 1.25$ mA/V^2

D 7.102 Using the feedback bias arrangement shown in Fig. 7.50 with a 5-V supply and an NMOS device for which $V_t = 1$ V and $k_n = 10$ mA/V^2, find R_D to establish a drain current of 0.2 mA.

D 7.103 Figure P7.103 shows a variation of the feedback-bias circuit of Fig. 7.50. Using a 5-V supply with an NMOS transistor for which $V_t = 0.8$ V, $k_n = 8$ mA/V^2, and $\lambda = 0$, provide a design that biases the transistor at $I_D = 1$ mA, with V_{DS} large enough to allow saturation operation for a 2-V negative signal swing at the drain. Use 22 MΩ as the largest resistor in the feedback-bias network. What values of R_D, R_{G1}, and R_{G2} have you chosen? Specify all resistors to two significant digits.

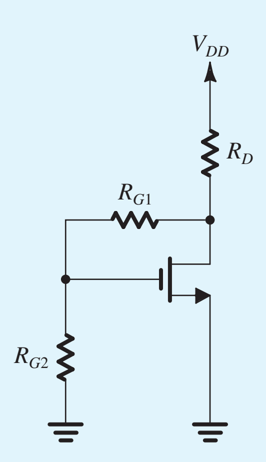


Figure P7.103

D 7.104 For the circuit in Fig. 7.51(a), neglect the base current I_B in comparison with the current in the voltage divider. It is required to bias the transistor at $I_C = 1$ mA, which requires selecting R_{B1} and R_{B2} so that $V_{BE} = 0.710$ V. If $V_{CC} = 3$ V, what must the ratio R_{B1}/R_{B2} be? Now, if R_{B1} and R_{B2} are 1% resistors, that is, each can be in the range of 0.99 to 1.01 of its nominal value, what is the range obtained for V_{BE}? What is the corresponding range of I_C? If $R_C = 2$ kΩ, what is the range obtained for V_{CE}? Comment on the efficacy of this biasing arrangement.

D 7.105 It is required to bias the transistor in the circuit of Fig. 7.51(b) at $I_C = 1$ mA. The transistor β is specified to be nominally 100, but it can fall in the range of 50 to 150. For $V_{CC} = +3$ V and $R_C = 2$ kΩ, find the required value of R_B to achieve $I_C = 1$ mA for the "nominal" transistor. What is the expected range for I_C and V_{CE}? Comment on the efficacy of this bias design.

D 7.106 Consider the single-supply bias network shown in Fig. 7.52(a). Provide a design using a 9-V supply in which the supply voltage is equally split between R_C, V_{CE}, and R_E with a collector current of 0.6 mA. The transistor β is specified to have a minimum value of 90. Use a voltage-divider current of $I_E/10$, or slightly higher. Since a reasonable design should operate for the best transistors for which β is very high, do your initial design with $\beta = \infty$. Then choose suitable 5% resistors (see Appendix J), making the choice in a way that will result in a V_{BB} that is slightly higher than the ideal value. Specify the values you have chosen for R_E, R_C, R_1, and R_2. Now, find V_B, V_E, V_C, and I_C for your final design using $\beta = 90$.

D 7.107 Repeat Problem 7.106, but use a voltage-divider current that is $I_E/2$. Check your design at $\beta = 90$. If you have the data available, find how low β can be while the value of I_C does not fall below that obtained with the design of Problem 7.106 for $\beta = 90$.

D *7.108 It is required to design the bias circuit of Fig. 7.52 for a BJT whose nominal $\beta = 100$.

(a) Find the largest ratio (R_B/R_E) that will guarantee I_E remains within ±5% of its nominal value for β as low as 50 and as high as 150.
(b) If the resistance ratio found in (a) is used, find an expression for the voltage $V_{BB} \equiv V_{CC}R_2/(R_1 + R_2)$ that will result in a voltage drop of $V_{CC}/3$ across R_E.
(c) For $V_{CC} = 5$ V, find the required values of R_1, R_2, and R_E to obtain $I_E = 0.5$ mA and to satisfy the requirement for stability of I_E in (a).
(d) Find R_C so that $V_{CE} = 1.0$ V for β equal to its nominal value.

Check your design by evaluating the resulting range of I_E.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

D *7.109 Consider the two-supply bias arrangement shown in Fig. 7.53 using ±5-V supplies. It is required to design the circuit so that $I_C = 0.5$ mA and V_C is placed 2 V above V_E.

(a) For $\beta = \infty$, what values of R_E and R_C are required?
(b) If the BJT is specified to have a minimum β of 50, find the largest value for R_B consistent with the need to limit the voltage drop across it to one-tenth the voltage drop across R_E.
(c) What standard 5% resistor values (see Appendix J) would you use for R_B, R_E, and R_C? In making your selection, use somewhat lower values in order to compensate for the low-β effects.
(d) For the values you selected in (c), find I_C, V_B, V_E, and V_C for $\beta = \infty$ and for $\beta = 50$.

D *7.110 Utilizing ±3-V power supplies, it is required to design a version of the circuit in Fig. 7.53 in which the signal will be coupled to the emitter and thus R_B can be set to zero. Find values for R_E and R_C so that a dc emitter current of 0.4 mA is obtained and so that the gain is maximized while allowing ±1 V of signal swing at the collector. If temperature increases from the nominal value of 25°C to 125°C, estimate the percentage change in collector bias current. In addition to the –2 mV/°C change in V_{BE}, assume that the transistor β changes over this temperature range from 50 to 150.

SIM D 7.111 Using a 3-V power supply, design a version of the circuit of Fig. 7.54 to provide a dc emitter current of 0.5 mA and to allow a ±1-V signal swing at the collector. The BJT has a nominal $\beta = 100$. Use standard 5% resistor values (see Appendix J). If the actual BJT used has $\beta = 50$, what emitter current is obtained? Also, what is the allowable signal swing at the collector? Repeat for $\beta = 150$.

D *7.112 (a) Using a 3-V power supply, design the feedback bias circuit of Fig. 7.54 to provide $I_C = 1$ mA and $V_C = V_{CC}/2$ for $\beta = 100$.
(b) Select standard 5% resistor values, and reevaluate V_C and I_C for $\beta = 100$.
(c) Find V_C and I_C for $\beta = \infty$.
(d) To improve the situation that obtains when high-β transistors are used, we have to arrange for an additional current to flow through R_B. This can be achieved by connecting a resistor between base and emitter, as shown in Fig. P7.112. Design this circuit for $\beta = 100$. Use a current through R_{B2} equal to the base current. Now, what values of V_C and I_C result with $\beta = \infty$?

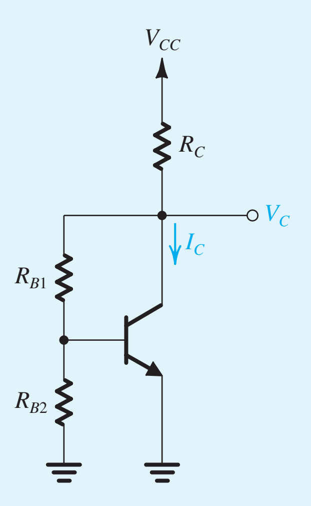


Figure P7.112

D 7.113 A circuit that can provide a very large voltage gain for a high-resistance load is shown in Fig. P7.113. Find the values of I and R_B to bias the BJT at $I_C = 1$ mA and $V_C = 1.5$ V. Let $\beta = 100$.

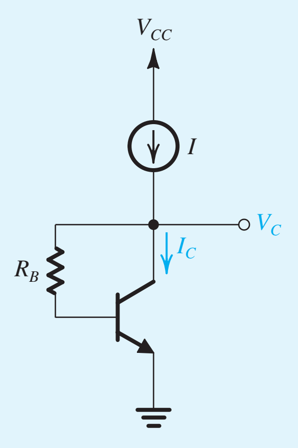


Figure P7.113

7.114 The circuit in Fig. P7.114 provides a constant current I_O as long as the circuit to which the collector is

connected maintains the BJT in the active mode. Show that

$$I_O = \alpha \frac{V_{CC}[R_2/(R_1+R_2)] - V_{BE}}{R_E + (R_1 \| R_2)/(\beta+1)}$$

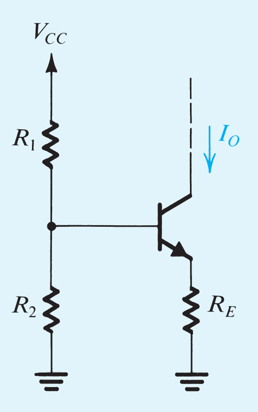


Figure P7.114

SIM D ***7.115** For the circuit in Fig. P7.115, assuming all transistors to be identical with β infinite, derive an expression for the output current I_O, and show that by selecting

$$R_1 = R_2$$

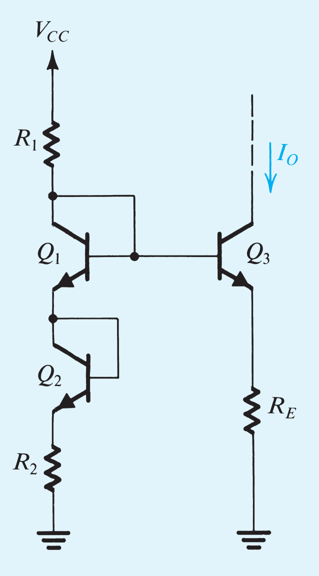


Figure P7.115

and keeping the current in each junction the same, the current I_O will be

$$I_O = \frac{V_{CC}}{2R_E}$$

which is independent of V_{BE}. What must the relationship of R_E to R_1 and R_2 be? For $V_{CC} = 10$ V and $V_{BE} = 0.7$ V, design the circuit to obtain an output current of 0.5 mA. What is the lowest voltage that can be applied to the collector of Q_3?

D **7.116** For the circuit in Fig. P7.116 find the value of R that will result in $I_O \simeq 0.5$ mA. What is the largest voltage that can be applied to the collector? Assume $|V_{BE}| = 0.7$ V.

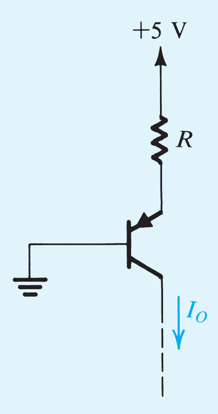


Figure P7.116

Section 7.5: Discrete-Circuit Amplifiers

7.117 Calculate the overall voltage gain G_v of a common-source amplifier for which g_m = 3 mA/V, r_o = 100 kΩ, R_D = 10 kΩ, and R_G = 10 MΩ. The amplifier is fed from a signal source with a Thévenin resistance of 1 MΩ, and the amplifier output is coupled to a load resistance of 20 kΩ.

SIM **7.118** The NMOS transistor in the CS amplifier shown in Fig. P7.118 has $V_t = 0.7$ V and $V_A = 50$ V.

(a) Neglecting the Early effect, verify that the MOSFET is operating in saturation with $I_D = 0.5$ mA and $V_{OV} = 0.3$ V. What must the MOSFET's k_n be? What is the dc voltage at the drain?
(b) Find R_{in} and G_v.
(c) If v_{sig} is a sinusoid with a peak amplitude $\hat{v}_{sig}$, find the maximum allowable value of $\hat{v}_{sig}$ for which the transistor remains in saturation. What is the corresponding amplitude of the output voltage?

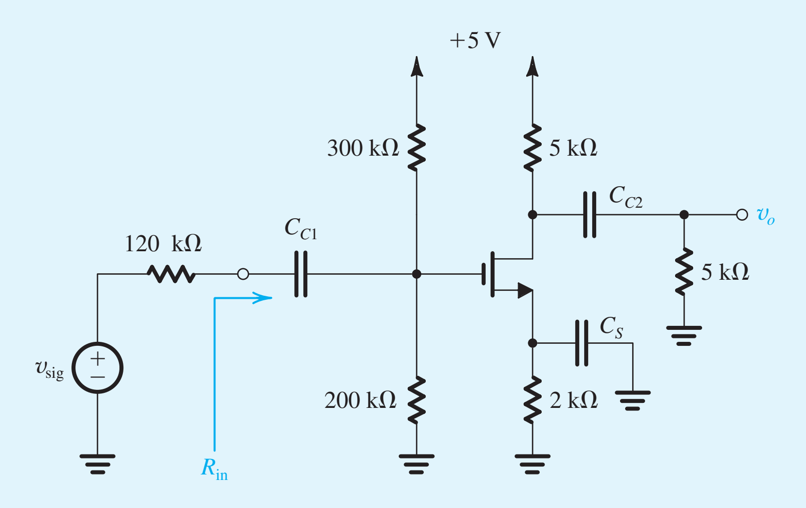


Figure P7.118

(d) What is the value of resistance R_s that needs to be inserted in series with capacitor C_S in order to allow us to double the input signal $\hat{v}_{sig}$? What output voltage now results?

SIM D *7.119 The PMOS transistor in the CS amplifier of Fig. P7.119 has $V_{tp} = -0.7$ V and a very large $|V_A|$.

(a) Select a value for R_S to bias the transistor at $I_D = 0.3$ mA and $|V_{OV}| = 0.3$ V. Assume v_{sig} to have a zero dc component.
(b) Select a value for R_D that results in $G_v = -10$ V/V.
(c) Find the largest sinusoid $\hat{v}_{sig}$ that the amplifier can handle while remaining in the saturation region. What is the corresponding signal at the output?
(d) If to obtain reasonably linear operation, $\hat{v}_{sig}$ is limited to 50 mV, what value can R_D be increased to while maintaining saturation-region operation? What is the new value of G_v?

7.120 Figure P7.120 shows a scheme for coupling and amplifying a high-frequency pulse signal. The circuit utilizes

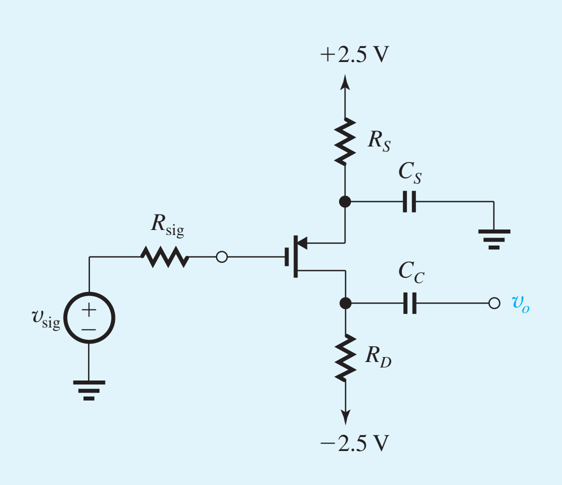


Figure P7.119

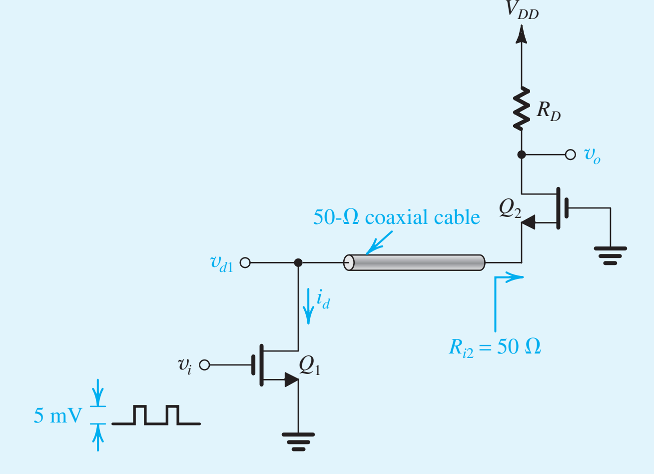
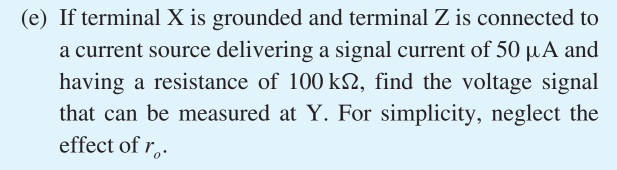


Figure P7.120

two MOSFETs whose bias details are not shown and a 50-Ω coaxial cable. Transistor Q_1 operates as a CS amplifier and Q_2 as a CG amplifier. For proper operation, transistor Q_2 is required to present a 50-Ω resistance to the cable. This situation is known as “proper termination” of the cable and ensures that there will be no signal reflection coming back on the cable. When the cable is properly terminated, its input resistance is 50 Ω. What must g_{m2} be? If Q_1 is biased at the same point as Q_2, what is the amplitude of the current pulses in the drain of Q_1? What is the amplitude of the voltage pulses at the drain of Q_1? What value of R_D is required to provide 1-V pulses at the drain of Q_2?

D *7.121 The MOSFET in the circuit of Fig. P7.121 has $V_t = 0.8$ V, $k_n = 5$ mA/V^2, and $V_A = 40$ V.

(a) Find the values of R_S, R_D, and R_G so that $I_D = 0.4$ mA, the largest possible value for R_D is used while a maximum signal swing at the drain of ±0.8 V is possible, and the input resistance at the gate is 10 MΩ. Neglect the Early effect.
(b) Find the values of g_m and r_o at the bias point.
(c) If terminal Z is grounded, terminal X is connected to a signal source having a resistance of 1 MΩ, and terminal Y is connected to a load resistance of 10 kΩ, find the voltage gain from signal source to load.
(d) If terminal Y is grounded, find the voltage gain from X to Z with Z open-circuited. What is the output resistance of the source follower?



(e) If terminal X is grounded and terminal Z is connected to a current source delivering a signal current of 50 μA and having a resistance of 100 kΩ, find the voltage signal that can be measured at Y. For simplicity, neglect the effect of r_o.

***7.122**

(a) The NMOS transistor in the source-follower circuit of Fig. P7.122(a) has $g_m = 10$ mA/V and a large r_o. Find the open-circuit voltage gain and the output resistance.
(b) The NMOS transistor in the common-gate amplifier of Fig. P7.122(b) has $g_m = 10$ mA/V and a large r_o. Find the input resistance and the voltage gain.
(c) If the output of the source follower in (a) is connected to the input of the common-gate amplifier in (b), use the results of (a) and (b) to obtain the overall voltage gain v_o/v_i.

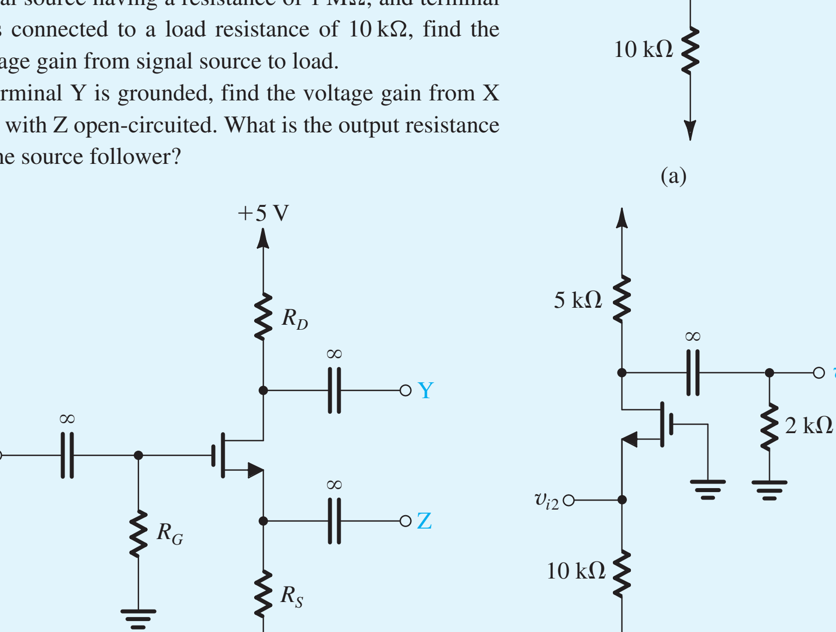


Figure P7.121

Figure P7.122

D **7.123 The MOSFET in the amplifier circuit of Fig. P7.123 has $V_t = 0.6$ V, $k_n = 5$ mA/V^2, and $V_A = 60$ V. The signal v_{sig} has a zero average.

(a) It is required to bias the transistor to operate at an overdrive voltage $V_{OV} = 0.2$ V. What must the dc voltage at the drain be? Calculate the dc drain current I_D taking into account V_A. Now, what value must the drain resistance R_D have?

(b) Calculate the values of g_m and r_o at the bias point established in (a).

(c) Using the small-signal equivalent circuit of the amplifier, show that the voltage gain is given by

$$\frac{v_o}{v_{sig}} = -\frac{R_2/R_1}{1 + \dfrac{1 + R_2/R_1}{g_m(R_D \| r_o \| R_2)(1 - 1/g_m R_2)}}$$

and find the value of the gain.

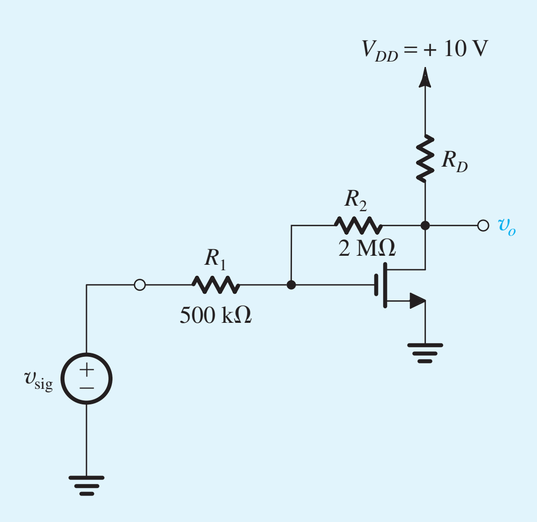


Figure P7.123

P.S. This feedback amplifier and the gain expression should remind you of an op amp utilized in the inverting configuration. We shall study feedback formally in Chapter 11.

D **7.124 The MOSFET in the amplifier circuit of Fig. P7.124 has $V_t = 0.6$ V and $k_n = 5$ mA/V^2. We shall assume that V_A is sufficiently large so that we can ignore the Early effect. The input signal v_{sig} has a zero average.

(a) It is required to bias the transistor to operate at an overdrive voltage $V_{OV} = 0.2$ V. What must the dc voltage at the drain be? Calculate the dc drain current I_D. What value must R_D have?

(b) Calculate the value of g_m at the bias point.

(c) Use the small-signal equivalent circuit of the amplifier to show that

$$\frac{v_o}{v_{sig}} = \frac{1 + (R_2/R_1)}{1 + \dfrac{(1 + R_2/R_1)}{g_m R_D'}}$$

and

$$R_{in} = \frac{1}{g_m}\left(1 + g_m R_D' \frac{R_1}{R_1 + R_2}\right)$$

where

$$R_D' = R_D \| (R_1 + R_2)$$

(d) Evaluate v_o/v_{sig} and R_{in}.

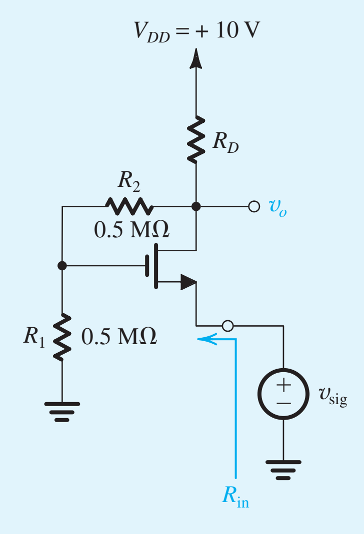


Figure P7.124

P.S. This feedback amplifier circuit and the gain formula should remind you of an op amp connected in the noninverting configuration. We shall study feedback formally in Chapter 11.

7.125 For the common-emitter amplifier shown in Fig. P7.125, let $V_{CC} = 15$ V, $R_1 = 27$ kΩ, $R_2 = 15$ kΩ, $R_E = 2.4$ kΩ, and $R_C = 3.9$ kΩ. The transistor has $\beta = 100$. Calculate the dc bias current I_C. If the amplifier operates between a source for which $R_{sig} = 2$ kΩ and a load of 2 kΩ, replace the transistor with its hybrid-π model, and find the values of R_{in}, and the overall voltage gain v_o/v_{sig}.

D 7.126 Using the topology of Fig. P7.125, design an amplifier to operate between a 2-kΩ source and a 2-kΩ load with a gain v_o/v_{sig} of –40 V/V. The power supply available is 15 V. Use an emitter current of approximately 2 mA and a current of about one-tenth of that in the voltage divider that feeds the base, with the dc voltage at the base about one-third of the supply. The transistor available has $\beta = 100$. Use standard 5% resistors (see Appendix J).

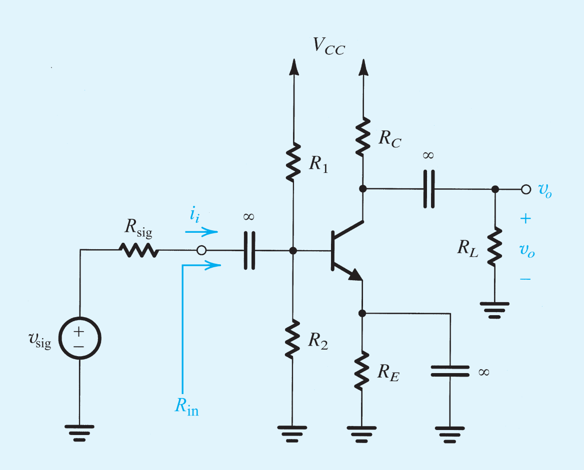


Figure P7.125

D 7.127 A designer, having examined the situation described in Problem 7.125 and estimating the available gain to be approximately –36.3 V/V, wants to explore the possibility of improvement by reducing the loading of the source by the amplifier input. As an experiment, the designer varies the resistance levels by a factor of approximately 3: R_1 to 82 kΩ, R_2 to 47 kΩ, R_E to 7.2 kΩ, and R_C to 12 kΩ (standard values of 5%-tolerance resistors). With $V_{CC} = 15$ V, $R_{sig} = 2$ kΩ, $R_L = 2$ kΩ, and $\beta = 100$, what does the gain become? Comment.

D 7.128 The CE amplifier circuit of Fig. P7.128 is biased with a constant-current source I. It is required to design the circuit (i.e., find values for I, R_B, and R_C) to meet the following specifications:

(a) $R_{in} \simeq 10$ kΩ.
(b) The dc voltage drop across R_B is approximately 0.2 V.
(c) The open-circuit voltage gain from base to collector is the maximum possible, consistent with the requirement that the collector voltage never fall by more than approximately 0.4 V below the base voltage with the signal between base and emitter being as high as 5 mV.

Assume that v_{sig} is a sinusoidal source, the available supply $V_{CC} = 5$ V, and the transistor has $\beta = 100$. Use standard 5% resistance values, and specify the value of I to one significant digit. What base-to-collector open-circuit voltage gain does your design provide? If $R_{sig} = R_L = 20$ kΩ, what is the overall voltage gain?

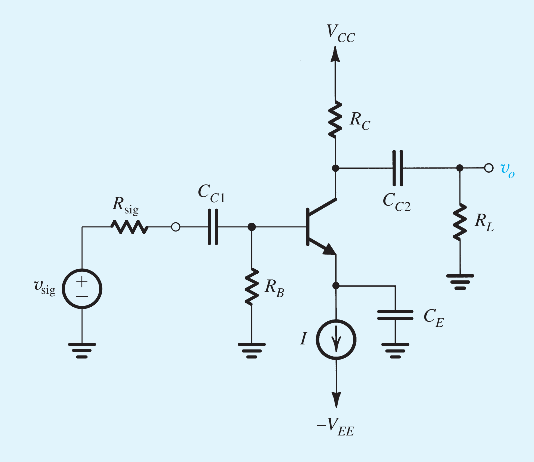


Figure P7.128

D 7.129 In the circuit of Fig. P7.129, v_{sig} is a small sine-wave signal with zero average. The transistor β is 100.

(a) Find the value of R_E to establish a dc emitter current of about 0.5 mA.

(b) Find R_C to establish a dc collector voltage of about +0.5 V.

(c) For $R_L = 10\ \text{k}\Omega$, draw the small-signal equivalent circuit of the amplifier and determine its overall voltage gain.

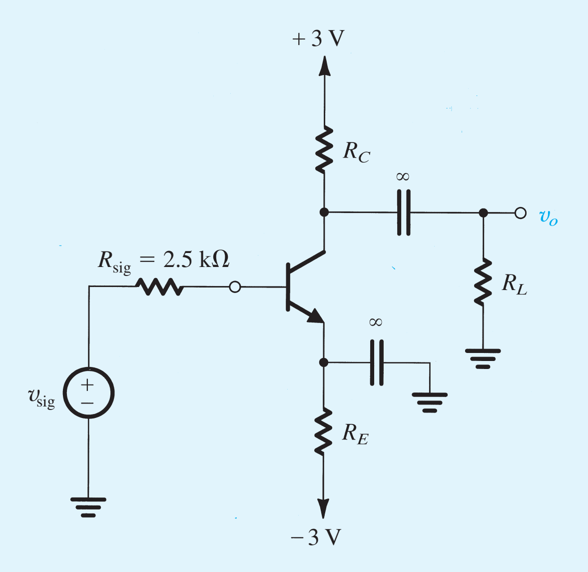


Figure P7.129

***7.130** The amplifier of Fig. P7.130 consists of two identical common-emitter amplifiers connected in cascade. Observe that the input resistance of the second stage, R_{in2}, constitutes the load resistance of the first stage.

(a) For $V_{CC} = 15$ V, $R_1 = 100\ \text{k}\Omega$, $R_2 = 47\ \text{k}\Omega$, $R_E = 3.9\ \text{k}\Omega$, $R_C = 6.8\ \text{k}\Omega$, and $\beta = 100$, determine the dc collector current and dc collector voltage of each transistor.

(b) Draw the small-signal equivalent circuit of the entire amplifier and give the values of all its components.

(c) Find R_{in1} and v_{b1}/v_{sig} for $R_{sig} = 5\ \text{k}\Omega$.

(d) Find R_{in2} and v_{b2}/v_{b1}.

(e) For $R_L = 2\ \text{k}\Omega$, find v_o/v_{b2}.

(f) Find the overall voltage gain v_o/v_{sig}.

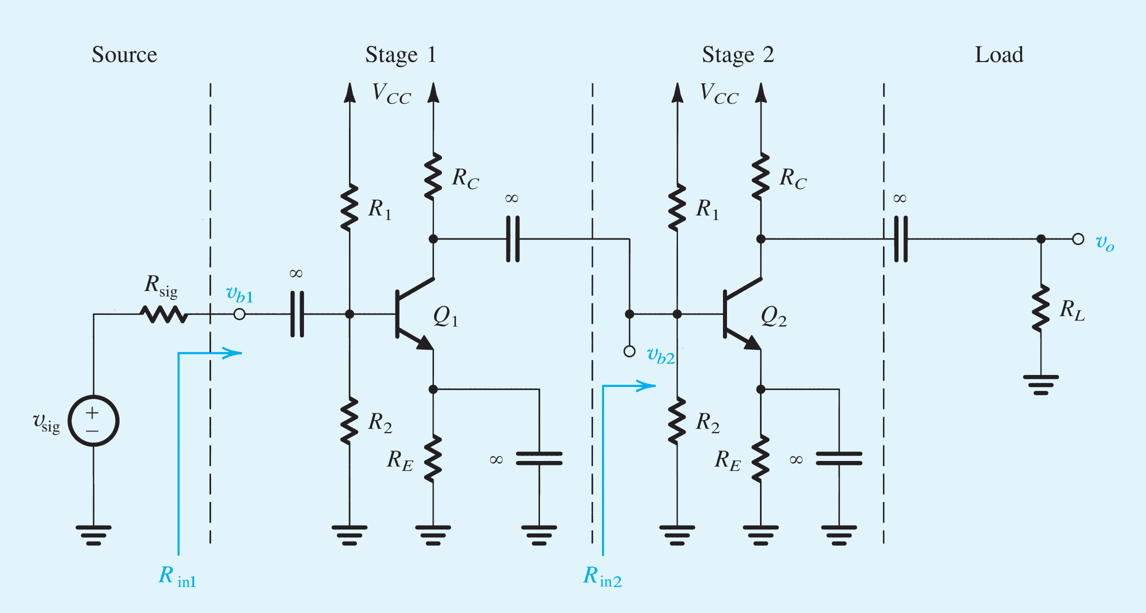


Figure P7.130

7.131 In the circuit of Fig. P7.131, the BJT is biased with a constant-current source, and v_{sig} is a small sine-wave signal. Find R_{in} and the gain v_o/v_{sig}. Assume $\beta = 100$. If the amplitude of the signal v_{be} is to be limited to 5 mV, what is the largest signal at the input? What is the corresponding signal at the output?

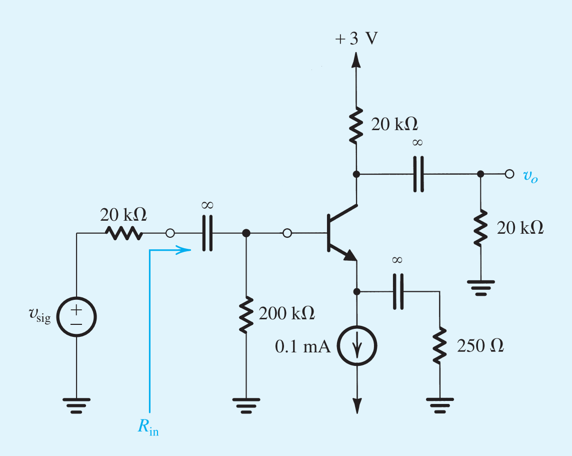


Figure P7.131

***7.132** The BJT in the circuit of Fig. P7.132 has $\beta = 100$.

(a) Find the dc collector current and the dc voltage at the collector.

(b) Replacing the transistor by its T model, draw the small-signal equivalent circuit of the amplifier. Analyze the resulting circuit to determine the voltage gain v_o/v_i.

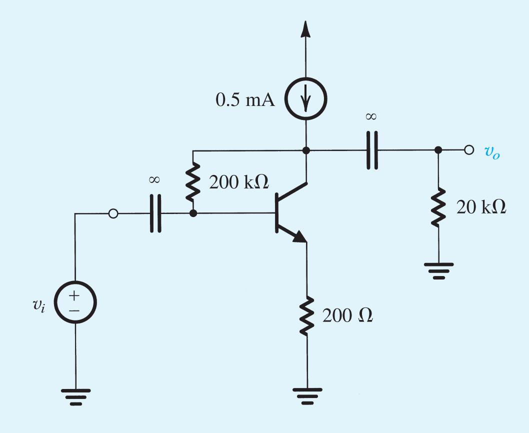
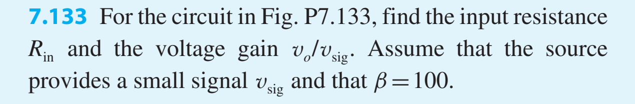


Figure P7.132



7.133 For the circuit in Fig. P7.133, find the input resistance R_{in} and the voltage gain v_o/v_{sig}. Assume that the source provides a small signal v_{sig} and that $\beta = 100$.

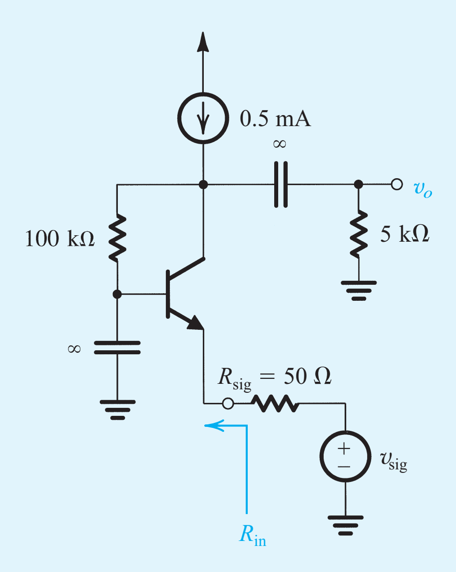


Figure P7.133

7.134 For the emitter-follower circuit shown in Fig. P7.134, the BJT used is specified to have β values in the range of 50 to 200 (a distressing situation for the circuit designer).

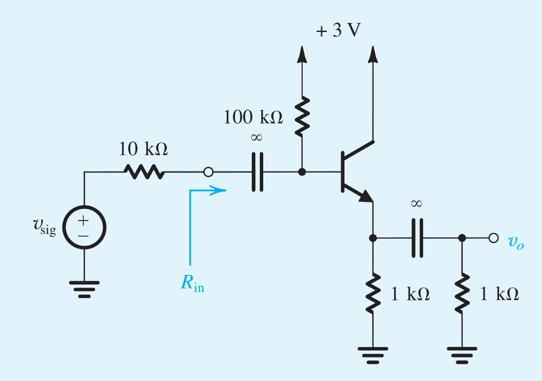


Figure P7.134

For the two extreme values of β ($\beta = 50$ and $\beta = 200$), find:

(a) I_E, V_E, and V_B
(b) the input resistance R_{in}
(c) the voltage gain v_o/v_{sig}

7.135 For the emitter follower in Fig. P7.135, the signal source is directly coupled to the transistor base. If the dc component of v_{sig} is zero, find the dc emitter current. Assume $\beta = 100$. Neglecting r_o, find R_{in}, the voltage gain v_o/v_{sig}, the current gain i_o/i_i, and the output resistance R_{out}.

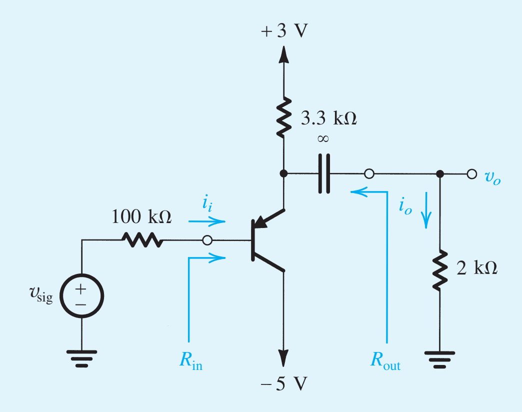


Figure P7.135

****7.136** For the circuit in Fig. P7.136, called a **bootstrapped follower:**

(a) Find the dc emitter current and g_m, r_e, and r_π. Use $\beta = 100$.
(b) Replace the BJT with its T model (neglecting r_o), and analyze the circuit to determine the input resistance R_{in} and the voltage gain v_o/v_{sig}.
(c) Repeat (b) for the case when capacitor C_B is open-circuited. Compare the results with those obtained in (b) to find the advantages of bootstrapping.

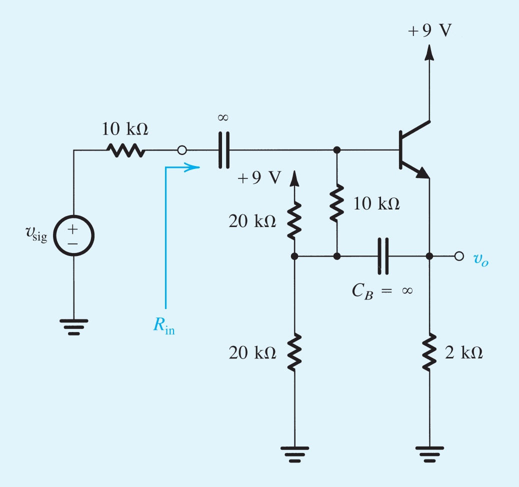


Figure P7.136

****7.137** For the follower circuit in Fig. P7.137, let transistor Q_1 have $\beta = 50$ and transistor Q_2 have $\beta = 100$, and neglect the effect of r_o. Use $V_{BE} = 0.7$ V.

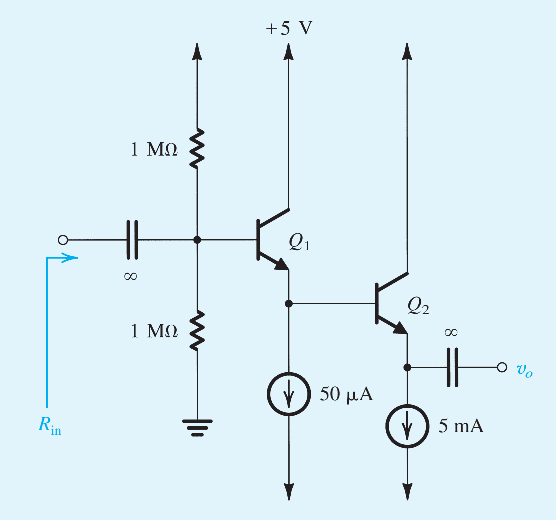


Figure P7.137

(a) Find the dc emitter currents of Q_1 and Q_2. Also, find the dc voltages V_{B1} and V_{B2}.

(b) If a load resistance $R_L = 1\ \text{k}\Omega$ is connected to the output terminal, find the voltage gain from the base to the emitter of Q_2, v_o/v_{b2}, and find the input resistance R_{ib2} looking into the base of Q_2. (*Hint:* Consider Q_2 as an emitter follower fed by a voltage v_{b2} at its base.)

(c) Replacing Q_2 with its input resistance R_{ib2} found in (b), analyze the circuit of emitter follower Q_1 to determine its input resistance R_{in}, and the gain from its base to its emitter, v_{e1}/v_{b1}.

(d) If the circuit is fed with a source having a 100-kΩ resistance, find the transmission to the base of Q_1, v_{b1}/v_{sig}.

(e) Find the overall voltage gain v_o/v_{sig}.

D 7.138 A CE amplifier has a midband voltage gain of $|A_M| = 100$ V/V, a lower 3-dB frequency of $f_L = 100$ Hz, and a higher 3-dB frequency $f_H = 500$ kHz. In Chapter 10 we will learn that connecting a resistance R_e in the emitter of the BJT results in lowering f_L and raising f_H by the factor $(1 + g_m R_e)$. If the BJT is biased at $I_C = 1$ mA, find R_e that will result in f_H at least equal to 2 MHz. What will the new values of f_L and A_M be?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

PART II

Integrated-Circuit Amplifiers

Having studied the MOSFET and the BJT and become familiar with their basic circuit applications, we are now ready to consider their use in the design of practical amplifier circuits that can be fabricated in integrated-circuit (IC) form. Part II is devoted to this rich subject. Its six chapters constitute a coherent treatment of IC amplifier design and can thus serve as a second course in electronic circuits.

Beginning with a brief introduction to the philosophy of IC design, Chapter 8 presents the basic circuit building blocks that are utilized in the design of IC amplifiers. However, the most important building block of all, the differential-pair configuration, is deferred to Chapter 9, where it is the main topic. Chapter 9 also considers the design of amplifiers that require a number of cascaded stages.

As mentioned at various points in Part I, amplifiers have finite bandwidths. Chapter 10 is devoted to the frequency-response analysis of amplifiers; it provides a comprehensive study of the mechanisms that limit the bandwidth and the tools and methods that are utilized to estimate it for a wide variety of amplifier circuit configurations. While the study of the first half or so of Chapter 10 is essential, some of its later sections can be postponed until a later point in the course or even to subsequent courses.

An essential tool in amplifier design is the judicious use of feedback. Chapter 11 deals with this exceedingly important subject. A thorough understanding of feedback concepts, insight into feedback configurations, and proficiency in the use of the feedback-analysis method are invaluable to the serious circuit designer.

In Chapter 12, we switch gears from dealing with primarily small-signal amplifiers to studying those that are required to handle large signals and large amounts of power. Finally, Chapter 13 brings together all the topics of Part II in an important application: namely, the design of operational-amplifier circuits. We will then have come full circle, from considering the op amp as a black box in Chapter 2 to understanding what is inside the box in Chapter 13.

Throughout Part II, MOSFET and BJT circuits are treated side by side. Because over 90% of ICs today employ the MOSFET, its circuits are presented first. Nevertheless, BJT circuits are presented with equal depth, although sometimes somewhat more briefly. In this regard, we draw the reader’s attention to Appendix G (on the website), which presents a valuable compilation of the properties of transistors of both types, allowing interesting comparisons to be made. As well, typical device parameter values are provided in Appendix K for a number of CMOS and bipolar fabrication process technologies.

CHAPTER 8

Building Blocks of Integrated-Circuit Amplifiers

IN THIS CHAPTER YOU WILL LEARN

1. The basic integrated-circuit (IC) design philosophy and how it differs from that for discrete-circuit design.
2. How current sources are used to bias IC amplifiers and how the use of current mirrors allows the replication of the reference current generated in one location at various other locations on the IC chip.
3. The basic gain cells of IC amplifiers, namely, the CS and CE amplifiers with current-source loads.
4. How the CG and CB amplifiers act as current buffers.
5. How to increase the gain realized in the basic gain cells by employing the principle of cascoding.
6. Analysis and design of the cascode amplifier and the cascode current source in both their MOS and bipolar forms.
7. Some ingenious analog circuit-design techniques that result in current mirrors with vastly improved characteristics.
8. How to pair transistors to realize amplifiers with characteristics superior to those obtained from a single-transistor stage.

Introduction

Having studied the two major transistor types, the MOSFET and the BJT, and their basic discrete-circuit amplifier configurations, we are now ready to begin the study of integrated-circuit (IC) amplifiers. This chapter is devoted to the design of the basic building blocks of IC amplifiers.

We begin with a brief section on the design philosophy of integrated circuits and how it differs from that of discrete circuits. This is followed by the study of IC biasing in Section 8.2, highlighting the design of current sources and current mirrors. The current mirror is one of the most important building blocks of analog integrated circuits. More advanced mirror circuits are presented in Section 8.6.

The heart of this chapter is the material in Sections 8.3 to 8.5. In Section 8.3 we present the basic gain cell of IC amplifiers, namely, the current-source-loaded common-source (common-emitter) amplifier. Then, in determining how to increase its gain, we discover the

need for current buffers. The two amplifier configurations capable of implementing a current buffer, the common-gate and common-base amplifiers, are studied in Section 8.4. This study differs from that in Chapter 7 in that r_o of the transistor is included, as must always be the case in integrated circuits. The study of the CG and CB leads naturally and seamlessly to the principle of cascoding and its applications in amplifier design: namely, the cascode amplifier and the cascode current source, which are very important building blocks of IC amplifiers.

The chapter concludes with the presentation in Section 8.7 of an interesting and useful collection of amplifier configurations, each utilizing a pair of transistors. Throughout this chapter, MOS and bipolar circuits are presented side by side, which allows a certain economy in presentation and, more important, provides an opportunity to compare and contrast the two circuit types.

8.1 IC Design Philosophy

Integrated-circuit fabrication technology (Appendix A) imposes constraints on and provides opportunities to the circuit designer. To cope with the constraints and take advantage of the opportunities, IC designers have over the years invented (and continue to invent) many ingenious techniques, and a distinct philosophy has emerged for the design of integrated circuits. In the following we provide a brief summary of the important constraints and opportunities and the major features of the IC design philosophy.

1. **Resistors.** To minimize the chip area, large and even moderate-size resistors are to be avoided. As well, economic considerations discourage the use of resistors of precise values. On the other hand, transistors can be made small and cheaply, and the designer is encouraged to use transistors in preference to resistors wherever possible. As a result, the classical biasing arrangement, popular in discrete-circuit amplifier design, is abandoned in IC amplifiers in favor of biasing with constant-current sources implemented with transistors operating in the active mode. As well, the collector and drain resistors in amplifiers are replaced with constant-current sources that have much higher incremental resistance, thus providing larger gains.
2. **Capacitors.** Chip-area considerations also make it impossible to fabricate large-valued capacitors such as those employed for signal coupling and bypass in discrete-circuit amplifiers. As a result, IC amplifiers are all direct coupled and utilize clever techniques, which we will study in this chapter and the next.

   Small-size capacitors, in the picofarad and fraction-of-a-picofarad range, are easy to fabricate in IC MOS technology. Such capacitors can be combined with MOS amplifiers and MOS switches to realize a wide variety of signal-processing functions, both analog (Chapter 17) and digital (Chapter 15).
3. **Power Supplies.** To pack a large number of devices on the same IC chip, and thus reduce system cost and increase reliability, the trend has been to reduce the device dimensions. (For a discussion of Moore's law and device scaling, see Section 14.5.) By 2014, CMOS process technologies capable of producing devices with a 14-nm channel length were in use. To avoid breaking down the thin oxide layers (less than 1 nm) used in these devices, power supplies are limited to 1 V or so. Low power-supply voltages help with another major design challenge; namely, keeping the power dissipated in the chip within acceptable limits. However, the use of such low dc power-supply voltages presents the circuit designer with a host of challenges. For instance, MOS

transistors must be operated with overdrive voltages of only 0.1 V to 0.2 V. In our study of MOS amplifiers, we will frequently comment on such issues.

4. **Device Variety.** Unlike the designer of discrete circuits, who is limited to available off-the-shelf transistors, the IC designer has the freedom to specify the device dimensions and to utilize device matching and arrays of devices having dimensions with specified ratios. For instance, one can utilize an array of bipolar transistors whose emitter–base–junction areas have binary-weighted ratios. CMOS technology provides even more flexibility, with the W and L values of MOS transistors selected to fit a very wide range of design requirements.
5. **Bipolar Technology.** BJTs are still used in special analog applications, such as high-quality general-purpose op-amp packages that are intended for assembly on printed-circuit (PC) boards (as opposed to being part of a system-on-a-chip). Bipolar circuits can also be combined with CMOS circuits in innovative and exciting ways in what is known as BiCMOS technology.
6. **CMOS Technology.** Currently the vast majority of analog integrated circuits are designed using CMOS technology. This practice was initially motivated by the need to be compatible with digital circuits, which have become predominantly CMOS. Now, however, the richness and the versatility that CMOS provides the analog designer is an even stronger reason for its dominance. We hope that the reader will come to appreciate this point in Chapters 8 and 9.

SOLID CIRCUITS WITH "FLYING WIRES":

As the importance of transistors grew during the 1950s, packaging became a problem. While it was possible to create smaller and smaller active devices, individual transistor packages had to be large enough to be held in the assembly of an electronic system. As one solution to this problem, Texas Instruments (TI) initiated a program of package modular electronics in the creation, on a ceramic substrate, of larger, more functional system elements. Employed to work in this direction, Jack Kilby believed it was necessary to go one step further and design the multiplicity of active and passive elements on a single piece of semiconductor. Thus, in 1958 he created the first "solid circuit," incorporating many transistors and resistors formed on a single slab of germanium and coupled by "flying wire" interconnections to form system elements such as oscillators and amplifiers. In 1959 TI began to use this technique to manufacture the 507 Binary Flip-Flop. While the approach was successful in producing small space-efficient modules, it was not suited for mass production. In 2000 Kilby received the Nobel Prize in Physics, in recognition of his part in the invention of the integrated circuit.

8.2 IC Biasing—Current Sources, Current Mirrors, and Current-Steering Circuits

Biasing in integrated-circuit design is based on the use of constant-current sources. On an IC chip with a number of amplifier stages, a constant dc current (called a **reference current**) is generated at one location and is then replicated at various other locations for biasing the various amplifier stages through a process known as **current steering**. This approach has the advantage that the effort expended on generating a predictable and stable reference current,

usually utilizing a precision resistor external to the chip or a special circuit on the chip, need not be repeated for every amplifier stage. Furthermore, the bias currents of the various stages track each other in case of changes in power-supply voltage or in temperature.

In this section we study circuit building blocks and techniques employed in the bias design of IC amplifiers. These current-source circuits are also utilized as amplifier load elements, as will be seen in Sections 8.3 and 8.4.

8.2.1 The Basic MOSFET Current Source

Figure 8.1 shows the circuit of a simple MOS constant-current source. The heart of the circuit is transistor Q_1, the drain of which is shorted to its gate,[1] thereby forcing it to operate in the saturation mode with

$$I_{D1} = \frac{1}{2}k_n'\left(\frac{W}{L}\right)_1 (V_{GS} - V_{tn})^2 \tag{8.1}$$

where we have neglected channel-length modulation. The drain current of Q_1 is supplied by V_{DD} through resistor R, which in most cases would be outside the IC chip. Since the gate currents are zero,

$$I_{D1} = I_{\text{REF}} = \frac{V_{DD} - V_{GS}}{R} \tag{8.2}$$

where the current through R is considered to be the reference current of the current source and is denoted I_{REF}. Equations (8.1) and (8.2) can be used to determine the value required for R.

Now consider transistor Q_2: It has the same V_{GS} as Q_1; thus, if we assume that it is operating in saturation, its drain current, which is the output current I_O of the current source, will be

$$I_O = I_{D2} = \frac{1}{2}k_n'\left(\frac{W}{L}\right)_2 (V_{GS} - V_{tn})^2 \tag{8.3}$$

where we have neglected channel-length modulation. Equations (8.1) and (8.3) enable us to relate the output current I_O to the reference current I_{REF} as follows:

$$\frac{I_O}{I_{\text{REF}}} = \frac{(W/L)_2}{(W/L)_1} \tag{8.4}$$

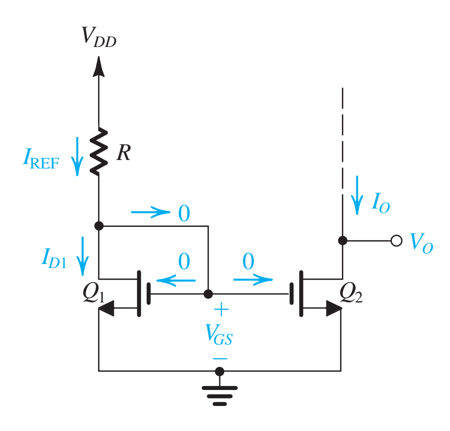


Figure 8.1 Circuit for a basic MOSFET constant-current source. For proper operation, the output terminal, that is, the drain of Q_2, must be connected to a circuit that ensures that Q_2 operates in saturation.

[1] Such a transistor is said to be *diode connected.*

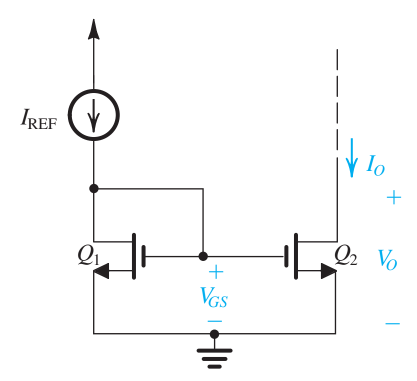


Figure 8.2 Basic MOSFET current mirror.

This is a simple and attractive relationship: The special connection of Q_1 and Q_2 provides an output current I_O that is related to the reference current I_{REF} by the aspect ratios of the transistors. In other words, the relationship between I_O and I_{REF} is solely determined by the geometries of the transistors. In the special case of identical transistors, $I_O = I_{REF}$, and the circuit simply replicates or mirrors the reference current in the output terminal. This has given the circuit composed of Q_1 and Q_2 the name **current mirror**, a name that is used irrespective of the ratio of device dimensions.

Figure 8.2 depicts the current-mirror circuit with the input reference current shown as being supplied by a current source for both simplicity and generality. The **current gain** or **current transfer ratio** of the current mirror is given by Eq. (8.4).

Effect of V_O on I_O In the description above for the operation of the current source of Fig. 8.1, we assumed Q_2 to be operating in saturation. This is essential if Q_2 is to supply a constant-current output. To ensure that Q_2 is saturated, the circuit to which the drain of Q_2 is to be connected must establish a drain voltage V_O that satisfies the relationship

$$V_O \geq V_{GS} - V_{tn} \tag{8.5}$$

or, equivalently, in terms of the overdrive voltage V_{OV} of Q_1 and Q_2,

$$V_O \geq V_{OV} \tag{8.6}$$

In other words, the current source will operate properly with an output voltage V_O as low as V_{OV}, which is a few tenths of a volt.

Although thus far neglected, channel-length modulation can have a significant effect on the operation of the current source. Consider, for simplicity, the case of identical devices Q_1 and Q_2. The drain current of Q_2, I_O, will equal the current in Q_1, I_{REF}, at the value of V_O that causes the two devices to have the same V_{DS}, that is, at $V_O = V_{GS}$. As V_O is increased above this value, I_O will increase according to the incremental output resistance r_{o2} of Q_2. This is illustrated in Fig. 8.3, which shows I_O versus V_O. Observe that since Q_2 is operating at a constant V_{GS} (determined by passing I_{REF} through the matched device Q_1), the curve in Fig. 8.3 is simply the i_D–v_{DS} characteristic curve of Q_2 for v_{GS} equal to the particular value V_{GS}.

In summary, the current source of Fig. 8.1 and the current mirror of Fig. 8.2 have a finite output resistance R_o,

$$R_o \equiv \frac{\Delta V_O}{\Delta I_O} = r_{o2} = \frac{V_{A2}}{I_O} \tag{8.7}$$

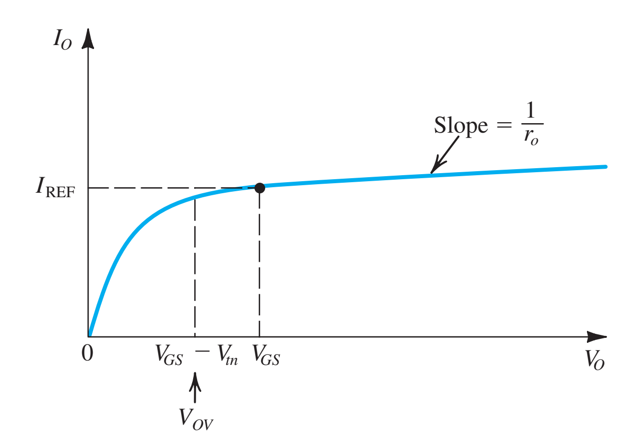


Figure 8.3 Output characteristic of the current source in Fig. 8.1 and the current mirror of Fig. 8.2 for the case of Q_2 matched to Q_1.

where I_O is given by Eq. (8.3) and V_{A2} is the Early voltage of Q_2. Also, recall that for a given process technology, V_A is proportional to the transistor channel length; thus, to obtain high output-resistance values, current sources are usually designed using transistors with relatively long channels. Finally, note that we can express the current I_O as

$$I_O = \frac{(W/L)_2}{(W/L)_1} I_{REF}\left(1 + \frac{V_O - V_{GS}}{V_{A2}}\right) \tag{8.8}$$

Example 8.1

Given $V_{DD} = 3$ V and using $I_{REF} = 100\ \mu$A, design the circuit of Fig. 8.1 to obtain an output current whose nominal value is 100 μA. Find R if Q_1 and Q_2 are matched and have channel lengths of 1 μm, channel widths of 10 μm, $V_t = 0.7$ V, and $k_n' = 200\ \mu\text{A/V}^2$. What is the lowest possible value of V_O? Assuming that for this process technology, the Early voltage $V_A' = 20$ V/μm, find the output resistance of the current source. Also, find the change in output current resulting from a +1-V change in V_O.

Solution

$$I_{D1} = I_{REF} = \frac{1}{2} k_n' \left(\frac{W}{L}\right)_1 V_{OV}^2$$

$$100 = \frac{1}{2} \times 200 \times 10\, V_{OV}^2$$

Thus,

$$V_{OV} = 0.316 \text{ V}$$

and

$$V_{GS} = V_t + V_{OV} = 0.7 + .316 \simeq 1 \text{ V}$$

$$R = \frac{V_{DD} - V_{GS}}{I_{REF}} = \frac{3 - 1}{0.1 \text{ mA}} = 20 \text{ k}\Omega$$

$$V_{Omin} = V_{OV} \simeq 0.3 \text{ V}$$

For the transistors used, $L = 1$ μm. Thus,

$$V_A = 20 \times 1 = 20 \text{ V}$$

$$r_{o2} = \frac{20 \text{ V}}{100 \text{ μA}} = 0.2 \text{ M}\Omega$$

The output current will be 100 μA at $V_O = V_{GS} = 1$ V. If V_O changes by +1 V, the corresponding change in I_O will be

$$\Delta I_O = \frac{\Delta V_O}{r_{o2}} = \frac{1 \text{ V}}{0.2 \text{ M}\Omega} = 5 \text{ μA}$$

EXERCISE

D8.1 In the current source of Example 8.1, it is required to reduce the change in output current, ΔI_O, corresponding to a change in output voltage, ΔV_O, of 1 V to 1% of I_O. What should the dimensions of Q_1 and Q_2 be changed to? Assume that Q_1 and Q_2 are to remain matched.
Ans. $L = 5$ μm; $W = 50$ μm

8.2.2 MOS Current-Steering Circuits

As mentioned earlier, once a constant current has been generated, it can be replicated to provide dc bias or load currents for the various amplifier stages in an IC. Current mirrors can obviously be used to implement this current-steering function. Figure 8.4 shows a simple current-steering circuit. Here Q_1 together with R determine the reference current I_{REF}. Transistors Q_1, Q_2, and Q_3 form a two-output current mirror,

$$I_2 = I_{REF} \frac{(W/L)_2}{(W/L)_1} \tag{8.9}$$

$$I_3 = I_{REF} \frac{(W/L)_3}{(W/L)_1} \tag{8.10}$$

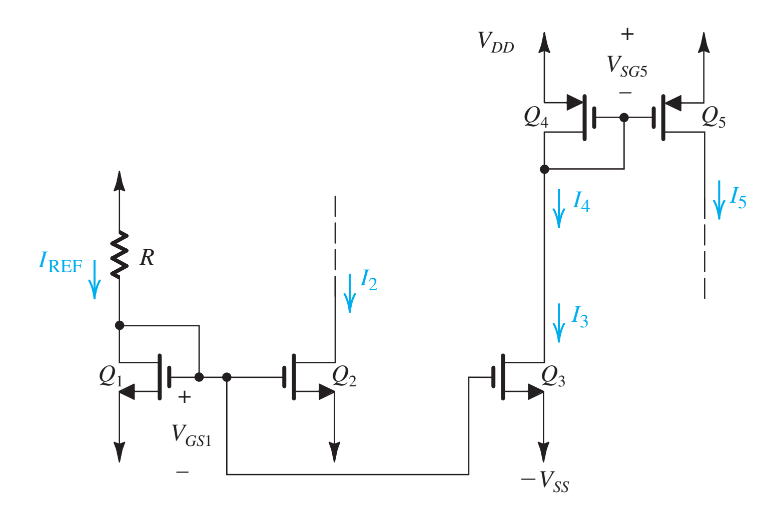


Figure 8.4 A current-steering circuit.

To ensure operation in the saturation region, the voltages at the drains of Q_2 and Q_3 are constrained as follows:

$$V_{D2}, V_{D3} \geq -V_{SS} + V_{GS1} - V_{tn} \tag{8.11}$$

or, equivalently,

$$V_{D2}, V_{D3} \geq -V_{SS} + V_{OV1} \tag{8.12}$$

where V_{OV1} is the overdrive voltage at which Q_1, Q_2, and Q_3 are operating. In other words, the drains of Q_2 and Q_3 will have to remain higher than $-V_{SS}$ by at least the overdrive voltage, which is usually a few tenths of a volt.

Continuing our discussion of the circuit in Fig. 8.4, we see that current I_3 is fed to the input side of a current mirror formed by PMOS transistors Q_4 and Q_5. This mirror provides

$$I_5 = I_4 \frac{(W/L)_5}{(W/L)_4} \tag{8.13}$$

where $I_4 = I_3$. To keep Q_5 in saturation, its drain voltage should be

$$V_{D5} \leq V_{DD} - |V_{OV5}| \tag{8.14}$$

where V_{OV5} is the overdrive voltage at which Q_5 is operating.

The constant current I_2 generated in the circuit of Fig. 8.4 can be used to bias a source-follower amplifier such as that implemented by transistor Q_6 in Fig. 8.5(a). Similarly, the constant current I_5 can be used as the load for a common-source amplifier such as that implemented with transistor Q_7 in Fig. 8.5(b). We will discuss the use of current sources as load elements for CS amplifiers in Section 8.3.

Finally, an important point to note is that in the circuit of Fig. 8.4, while Q_2 *pulls* its current I_2 from a circuit (not shown in Fig. 8.4), Q_5 *pushes* its current I_5 into a circuit (not shown in Fig. 8.4). Thus Q_5 is appropriately called a **current source**, whereas Q_2 should more properly be called a **current sink**. In an IC, both current sources and current sinks are usually needed. The difference between a current source and a current sink is further illustrated in Fig. 8.6, where V_{CSmin} denotes the minimum voltage needed across the current source (or sink) for its proper operation.

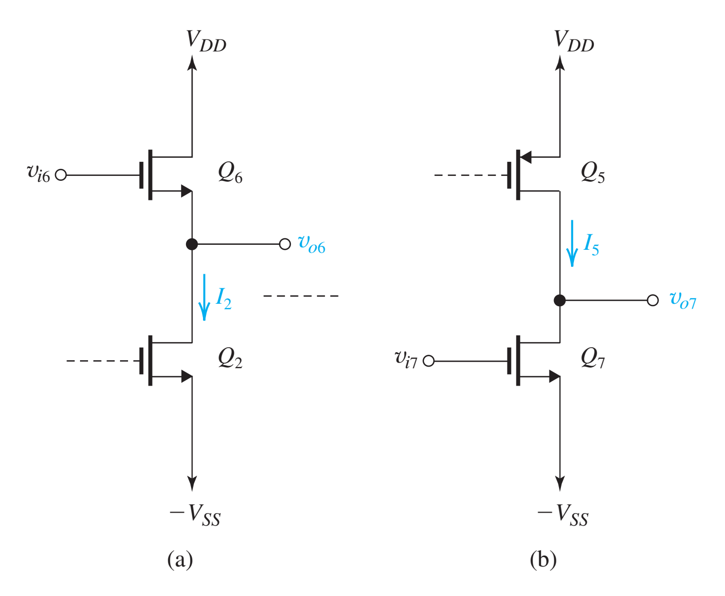


Figure 8.5 Application of the constant currents I_2 and I_5 generated in the current-steering circuit of Fig. 8.4. Constant-current I_2 is the bias current for the source follower Q_6, and constant-current I_5 is the load current for the common-source amplifier Q_7.

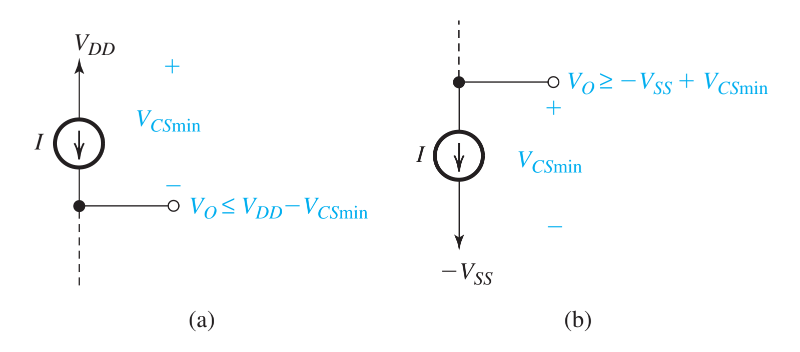


Figure 8.6 (**a**) A current source; and (**b**) a current sink.

EXERCISE

D8.2 For the circuit of Fig. 8.4, let $V_{DD} = V_{SS} = 1.5$ V, $V_{tn} = 0.6$ V, $V_{tp} = -0.6$ V, all channel lengths $= 1\ \mu$m, $k_n' = 200\ \mu\text{A/V}^2$, $k_p' = 80\ \mu\text{A/V}^2$, and $\lambda = 0$. For $I_{REF} = 10\ \mu$A, find the widths of all transistors to obtain $I_2 = 60\ \mu$A, $I_3 = 20\ \mu$A, and $I_5 = 80\ \mu$A. It is further required that the voltage at the drain of Q_2 be allowed to go down to within 0.2 V of the negative supply and that the voltage at the drain of Q_5 be allowed to go up to within 0.2 V of the positive supply.

Ans. $W_1 = 2.5\ \mu$m; $W_2 = 15\ \mu$m; $W_3 = 5\ \mu$m; $W_4 = 12.5\ \mu$m; $W_5 = 50\ \mu$m

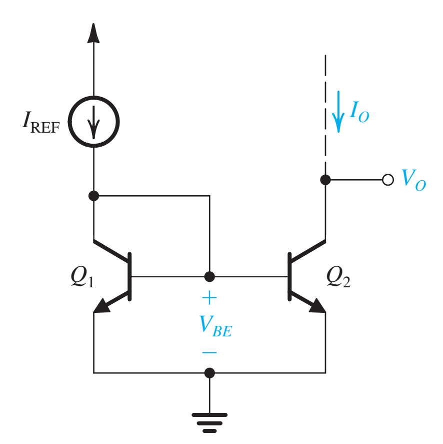


Figure 8.7 The basic BJT current mirror.

8.2.3 BJT Circuits

The basic BJT current mirror is shown in Fig. 8.7. It works in a fashion very similar to that of the MOS mirror. However, there are two important differences: First, the nonzero base current of the BJT (or, equivalently, the finite β) causes an error in the current transfer ratio of the bipolar mirror. Second, the current transfer ratio is determined by the relative areas of the emitter–base junctions of Q_1 and Q_2.

Let us first consider the case of β sufficiently high that we can neglect the base currents. The reference current I_{REF} is passed through the diode-connected transistor Q_1 and thus establishes a corresponding voltage V_{BE}, which in turn is applied between base and emitter of Q_2. Now, if Q_2 is matched to Q_1 or, more specifically, if the EBJ area of Q_2 is the same as that of Q_1, and thus Q_2 has the same scale current I_S as Q_1, then the collector current of Q_2 will be equal to that of Q_1; that is,

$$I_O = I_{REF} \tag{8.15}$$

For this to happen, however, Q_2 must be operating in the active mode, which in turn is achieved as long as the collector voltage V_O is 0.3 V or so higher than that of the emitter.

To obtain a current transfer ratio other than unity, say m, we simply arrange that the area of the EBJ of Q_2 is m times that of Q_1. In this case,

$$I_O = mI_{REF} \tag{8.16}$$

In general, the current transfer ratio is given by

$$\frac{I_O}{I_{REF}} = \frac{I_{S2}}{I_{S1}} = \frac{\text{Area of EBJ of } Q_2}{\text{Area of EBJ of } Q_1} \tag{8.17}$$

Alternatively, if the area ratio m is an integer, one can think of Q_2 as equivalent to m transistors, each matched to Q_1 and connected in parallel.

Next we consider the effect of finite transistor β on the current transfer ratio. The analysis for the case in which the current transfer ratio is nominally unity—that is, for the case in which Q_2 is matched to Q_1 —is illustrated in Fig. 8.8. The key point here is that since Q_1 and Q_2 are matched and have the same V_{BE}, their collector currents will be equal. The rest of the analysis is straightforward. A node equation at the collector of Q_1 yields

$$I_{REF} = I_C + 2I_C/\beta = I_C\left(1 + \frac{2}{\beta}\right)$$

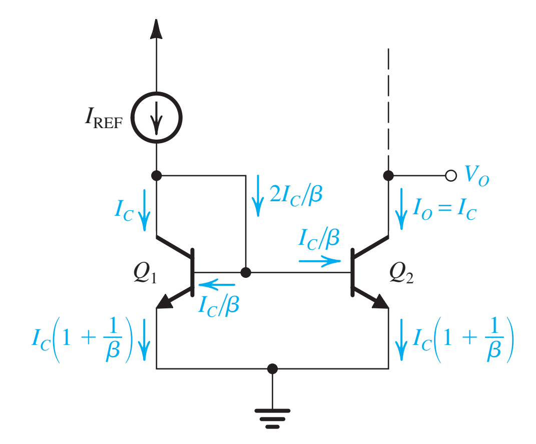


Figure 8.8 Analysis of the current mirror taking into account the finite β of the BJTs.

Finally, since $I_O = I_C$, the current transfer ratio can be found as

$$\frac{I_O}{I_{REF}} = \frac{I_C}{I_C\left(1+\dfrac{2}{\beta}\right)} = \frac{1}{1+\dfrac{2}{\beta}} \tag{8.18}$$

Note that as β approaches ∞, I_O/I_{REF} approaches the nominal value of unity. For typical values of β, however, the error in the current transfer ratio can be significant. For instance, $\beta = 100$ results in a 2% error in the current transfer ratio. Furthermore, the error due to the finite β increases as the nominal current transfer ratio is increased. The reader is encouraged to show that for a mirror with a nominal current transfer ratio m—that is, one in which $I_{S2} = mI_{S1}$—the actual current transfer ratio is given by

$$\frac{I_O}{I_{REF}} = \frac{m}{1+\dfrac{m+1}{\beta}} \tag{8.19}$$

In common with the MOS current mirror, the BJT mirror has a finite output resistance R_o,

$$R_o \equiv \frac{\Delta V_O}{\Delta I_O} = r_{o2} = \frac{V_{A2}}{I_O} \tag{8.20}$$

where V_{A2} and r_{o2} are the Early voltage and the output resistance, respectively, of Q_2. Thus, even if we neglect the error due to finite β, the output current I_O will be at its nominal value only when Q_2 has the same V_{CE} as Q_1, namely, at $V_O = V_{BE}$. As V_O is increased, I_O will correspondingly increase. Taking both the finite β and the finite R_o into account, we can express the output current of a BJT mirror with a nominal current transfer ratio m as

$$I_O = I_{REF}\frac{m}{1+\dfrac{m+1}{\beta}}\left(1+\frac{V_O - V_{BE}}{V_{A2}}\right) \tag{8.21}$$

where we note that the error term due to the Early effect is expressed in a form that shows that it reduces to zero for $V_O = V_{BE}$.

EXERCISE

8.3 Consider a BJT current mirror with a nominal current transfer ratio of unity. Let the transistors have $I_S = 10^{-15}$ A, $\beta = 100$, and $V_A = 100$ V. For $I_{\text{REF}} = 1$ mA, find I_O when $V_O = 5$ V. Also, find the output resistance.
Ans. 1.02 mA; 100 kΩ

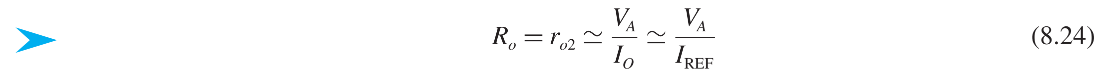
A Simple Current Source In a manner analogous to that in the MOS case, the basic BJT current mirror can be used to implement a simple current source, as shown in Fig. 8.9. Here the reference current is

$$I_{\text{REF}} = \frac{V_{CC} - V_{BE}}{R} \tag{8.22}$$

where V_{BE} is the base–emitter voltage corresponding to the desired value of I_{REF}. The output current I_O is given by

$$I_O = \frac{I_{\text{REF}}}{1 + (2/\beta)}\left(1 + \frac{V_O - V_{BE}}{V_A}\right) \tag{8.23}$$

The output resistance of this current source is r_o of Q_2,



$$R_o = r_{o2} \simeq \frac{V_A}{I_O} \simeq \frac{V_A}{I_{\text{REF}}} \tag{8.24}$$

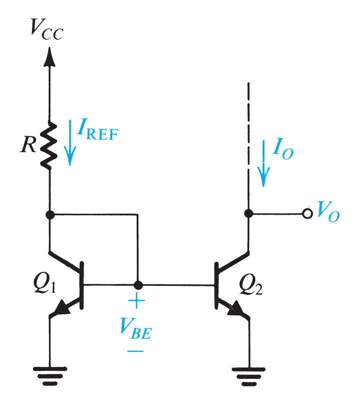


Figure 8.9 A simple BJT current source.

EXERCISE

D8.4 Assuming the availability of BJTs with scale currents $I_S = 10^{-15}$ A, $\beta = 100$, and $V_A = 50$ V, design the current-source circuit of Fig. 8.9 to provide an output current $I_O = 0.5$ mA at $V_O = 2$ V. The power supply $V_{CC} = 5$ V. Give the values of I_{REF}, R, and $V_{O\min}$. Also, find I_O at $V_O = 5$ V.
Ans. 0.497 mA; 8.71 kΩ; 0.3 V; 0.53 mA

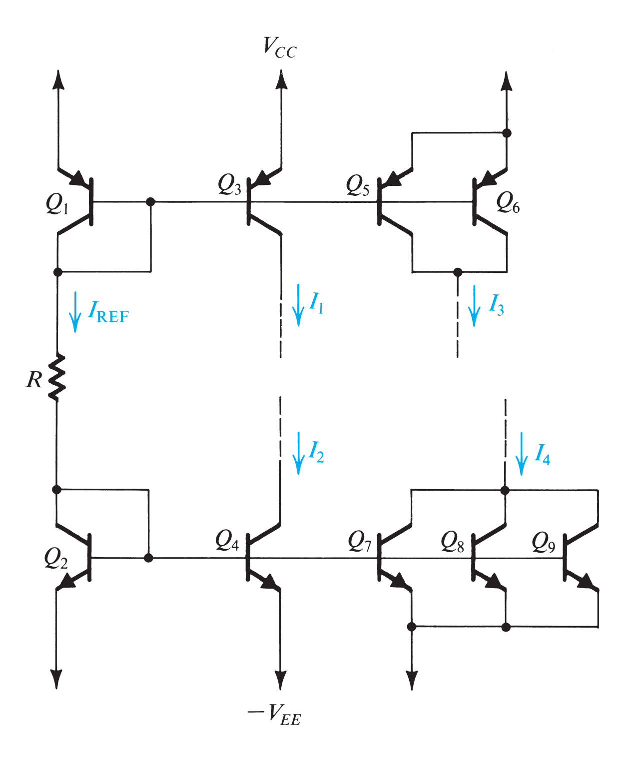


Figure 8.10 Generation of a number of constant currents of various magnitudes.

Current Steering To generate bias currents for different amplifier stages in an IC, the current-steering approach described for MOS circuits can be applied in the bipolar case. As an example, consider the circuit shown in Fig. 8.10. The dc reference current I_{REF} is generated in the branch that consists of the diode-connected transistor Q_1, resistor R, and the diode-connected transistor Q_2:

$$I_{REF} = \frac{V_{CC} + V_{EE} - V_{EB1} - V_{BE2}}{R} \tag{8.25}$$

Now, for simplicity, assume that all the transistors have high β and thus that the base currents are negligibly small. We will also neglect the Early effect. The diode-connected transistor Q_1 forms a current mirror with Q_3; thus Q_3 will supply a constant current I_1 equal to I_{REF}. Transistor Q_3 can supply this current to any load as long as the voltage that develops at the collector does not exceed ($V_{CC} - 0.3$ V); otherwise Q_3 would enter the saturation region.

To generate a dc current twice the value of I_{REF}, two transistors, Q_5 and Q_6, each of which is matched to Q_1, are connected in parallel, and the combination forms a mirror with Q_1. Thus $I_3 = 2I_{REF}$. Note that the parallel combination of Q_5 and Q_6 is equivalent to a transistor with an EBJ area double that of Q_1, which is precisely what is done when this circuit is fabricated in IC form.

Transistor Q_4 forms a mirror with Q_2; thus Q_4 provides a constant current I_2 equal to I_{REF}. Note that while Q_3 *sources* its current to parts of the circuit whose voltage should not exceed ($V_{CC} - 0.3$ V), Q_4 *sinks* its current from parts of the circuit whose voltage should not decrease below ($-V_{EE} + 0.3$ V). Finally, to generate a current three times I_{REF}, three transistors, Q_7, Q_8, and Q_9, each of which is matched to Q_2, are connected in parallel, and the combination is placed in a mirror configuration with Q_2. Again, in an IC implementation, Q_7, Q_8, and Q_9 would be replaced with a transistor having a junction area three times that of Q_2.

EXERCISE

8.5 Figure E8.5 shows an N-output current mirror. Assuming that all transistors are matched and have finite β and ignoring the effect of finite output resistances, show that

$$I_1 = I_2 = \cdots = I_N = \frac{I_{REF}}{1 + (N+1)/\beta}$$

For $\beta = 100$, find the maximum number of outputs for an error not exceeding 10%.

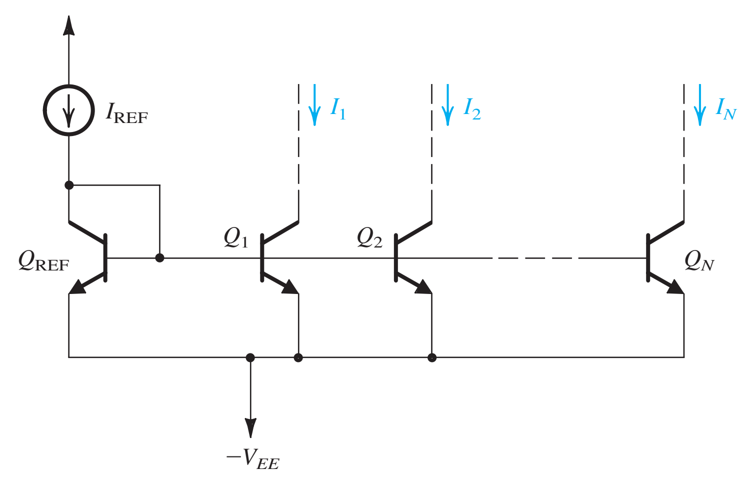


Figure E8.5

Ans. 9

A Bipolar Mirror with Base-Current Compensation Figure 8.11 shows a bipolar current mirror with a current transfer ratio that is much less dependent on β than that of the simple current mirror. The reduced dependence on β is achieved by including transistor Q_3, the emitter of which supplies the base currents of Q_1 and Q_2. The sum of the base currents is then divided by $(\beta_3 + 1)$, resulting in a much smaller error current that has to be supplied by I_{REF}. Detailed analysis is shown on the circuit diagram; it is based on the assumption that Q_1 and Q_2 are matched and thus have equal collector currents, I_C. A node equation at the node labeled x gives

$$I_{REF} = I_C\left[1 + \frac{2}{\beta(\beta+1)}\right]$$

Since

$$I_O = I_C$$

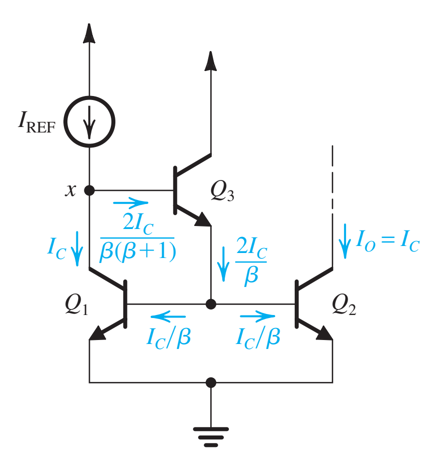


Figure 8.11 A current mirror with base-current compensation.

the current transfer ratio of the mirror will be

$$\frac{I_O}{I_{REF}} = \frac{1}{1 + 2/(\beta^2 + \beta)}$$
$$\simeq \frac{1}{1 + 2/\beta^2} \tag{8.26}$$

which means that the error due to finite β has been reduced from $2/\beta$ in the simple mirror to $2/\beta^2$, a tremendous improvement. Unfortunately, however, the output resistance remains approximately equal to that of the simple mirror, namely r_o. Finally, note that if a reference current I_{REF} is not available, we simply connect node x to the power supply, V_{CC}, through a resistance R. The result is a reference current given by

$$I_{REF} = \frac{V_{CC} - V_{BE1} - V_{BE3}}{R} \tag{8.27}$$

8.2.4 Small-Signal Operation of Current Mirrors

In addition to their use in biasing, current mirrors are sometimes employed as current amplifiers. It is therefore useful to derive the small-signal parameters of the current mirror, that is, R_{in}, A_{is}, and R_o.

Figure 8.12(a) shows a MOS current mirror biased with a dc input current I_{D1} and fed with a small-signal input current i_i. Note that V_{GS} and I_{D2} are the resulting dc quantities, while v_{gs} and i_o are signal quantities. Although we are not showing the circuit to which the output terminal is connected, we are assuming that the voltage at the drain of Q_2 exceeds the minimum required to keep Q_2 in saturation.

Replacing Q_1 and Q_2 with their small-signal models results in the circuit in Fig. 8.12(b). Observe that the controlled current source $g_{m1} v_{gs}$ appears across its control voltage v_{gs} and thus can be replaced by a resistance, $1/g_{m1}$, as shown in Fig. 8.12(c). For the latter circuit we can obtain

$$R_{in} = r_{o1} \left\| \frac{1}{g_{m1}} \simeq \frac{1}{g_{m1}} \right. \tag{8.28}$$

$$R_o = r_{o2} \tag{8.29}$$

$$A_{is} \equiv \left. \frac{i_o}{i_i} \right|_{v_{d2}=0} = \frac{g_{m2}\, v_{gs}}{i_i} \simeq \frac{g_{m2}\, i_i/g_{m1}}{i_i}$$

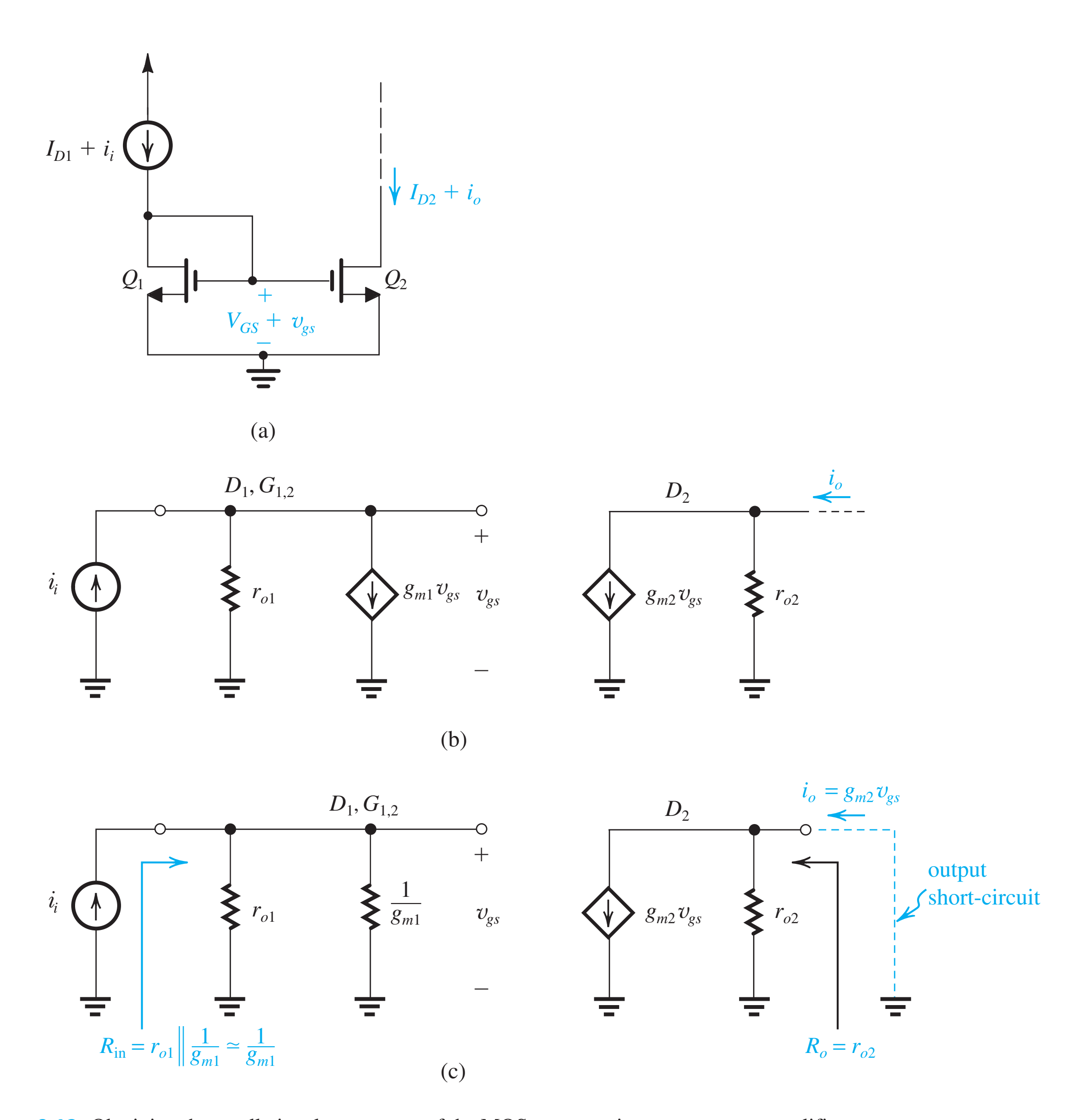


Figure 8.12 Obtaining the small-signal parameters of the MOS current mirror as a current amplifier.

Thus,

$$A_{is} = \frac{g_{m2}}{g_{m1}} \tag{8.30}$$

Substituting for $g_{m1,2} = \mu_n C_{ox}(W/L)_{1,2}V_{OV}$, where V_{OV} is the overdrive voltage at which Q_1 and Q_2 are operating, yields for the short-circuit current gain

$$A_{is} = \frac{(W/L)_2}{(W/L)_1} \tag{8.31}$$

which is equal to the dc or large-signal current transfer function—a clear indication of the excellent linearity of the current mirror.

We conclude that the current mirror is an excellent current amplifier: It has a relatively low input resistance ($1/g_{m1}$), a relatively high output resistance (r_{o2}), and a gain determined

by the aspect ratios of the MOSFETs. Finally, a similar development can be used to obtain the small-signal parameters of the bipolar mirror.

EXERCISE

D8.6 The MOSFETs in the current mirror of Fig. 8.12(a) have equal channel lengths, $\mu_n C_{ox} = 400\ \mu A/V^2$, and $V_A' = 20\ V/\mu m$. If the input bias current is 100 μA, find W_1, W_2, L_1, and L_2 to obtain a short-circuit current gain of 5, an input resistance of 1 kΩ, and an output resistance of 40 kΩ.
Ans. 12.5 μm; 62.5 μm; 1 μm; and 1 μm

THE INTEGRATED CIRCUIT:

In 1959, at the same time that Kilby and TI applied for a patent on "miniaturized electronic circuits," Robert Noyce (a cofounder of Fairchild Semiconductor and later of Intel) filed a patent on the "monolithic silicon-based integrated circuit." He later acknowledged the critical importance of Kurt Lehovac's idea of using reverse-biased junctions to isolate multiple devices on a single die. Lehovac, of Sprague Electric Company, also filed a patent in 1959. Regrettably, Noyce, who died in 1990, did not live to share in the Nobel Prize with Kilby.

8.3 The Basic Gain Cell

8.3.1 The CS and CE Amplifiers with Current-Source Loads

The basic gain cell in an IC amplifier is a common-source (CS) or common-emitter (CE) transistor loaded with a constant-current source, as shown in Fig. 8.13(a) and (b). These circuits are similar to the CS and CE amplifiers studied in Section 7.3, except that here we have replaced the resistances R_D and R_C with constant-current sources. This is done for two reasons: First, as mentioned in Section 8.1, it is difficult in IC technology to implement resistances with reasonably precise values; rather, it is much easier to use current sources, which are implemented using transistors. Second, by using a constant-current source we are in effect operating the CS and CE amplifiers with a very high (ideally infinite) load resistance; thus we can obtain a much higher gain than if a finite R_D or R_C is used. This is particularly the case because, even if passive resistances were available, they would have very small values because the dc power supplies are now limited to only 1 V to 2 V. These voltages, however, do allow the use of current sources that have large output resistances. The circuits in Fig. 8.13(a) and (b) are said to be **current-source loaded** or **active loaded**.

Before we consider the small-signal analysis of the active-loaded CS and CE amplifiers, a word on their dc bias is in order. Obviously, in each circuit Q_1 is biased at $I_D = I$ and $I_C = I$. But what determines the dc voltages at the drain (collector) and at the gate (base)? Usually, these gain cells will be part of larger circuits in which negative feedback is utilized to fix the values of V_{DS} and V_{GS} (V_{CE} and V_{BE}). In the next chapter we will begin to see complete IC amplifiers including biasing. For the time being, however, we shall assume that the MOS transistor in

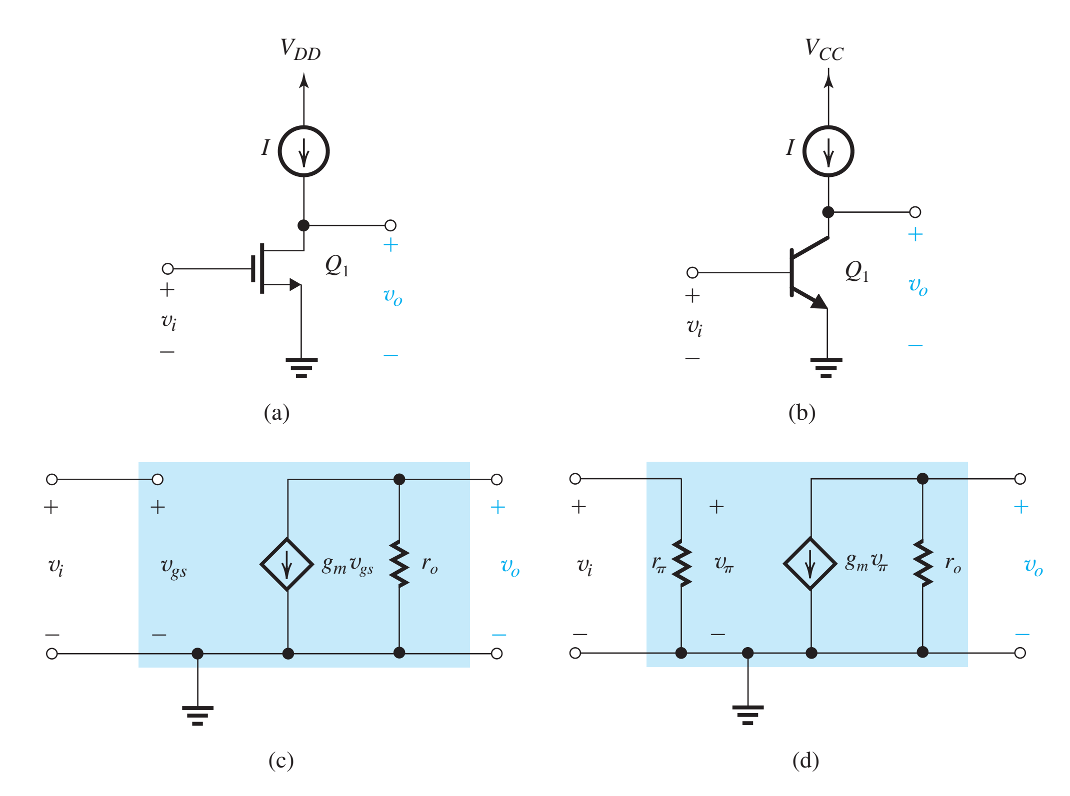


Figure 8.13 The basic gain cells of IC amplifiers: (**a**) current-source- or active-loaded common-source amplifier; (**b**) current-source- or active-loaded common-emitter amplifier; (**c**) small-signal equivalent circuit of (**a**); and (**d**) small-signal equivalent circuit of (**b**).

Fig. 8.13(a) is biased to operate in the saturation region and that the BJT in Fig. 8.13(b) is biased to operate in the active region. We will often refer to both the MOSFET and the BJT as operating in the "active region."

Small-signal analysis of the current-source-loaded CS and CE amplifiers can be performed by utilizing their equivalent-circuit models, shown respectively in Fig. 8.13(c) and (d). Observe that since the current-source load is assumed to be ideal, it is represented in the models by an infinite resistance. Practical current sources have finite output resistance, as we have seen in the previous section. For the time being, however, note that the CS and CE amplifiers of Fig. 8.13 are in effect operating in an open-circuit fashion. The only resistance between their output node and ground is the output resistance of the transistor itself, r_o. Thus the voltage gain obtained in these circuits is the maximum possible for a CS or a CE amplifier.

From Fig. 8.13(c) we obtain for the active-loaded CS amplifier:



$$R_{\text{in}} = \infty \tag{8.32}$$



$$A_{vo} = -g_m r_o \tag{8.33}$$

$$R_o = r_o \tag{8.34}$$

Similarly, from Fig. 8.13(d) we obtain for the active-loaded CE amplifier:

$$R_{in} = r_\pi \tag{8.35}$$

$$A_{vo} = -g_m r_o \tag{8.36}$$

$$R_o = r_o \tag{8.37}$$

Thus both circuits realize a voltage gain of magnitude $g_m r_o$. Since this is the maximum gain obtainable in a CS or CE amplifier, we refer to it as the **intrinsic gain** and give it the symbol A_0. Furthermore, it is useful to examine the nature of A_0 in a little more detail.

8.3.2 The Intrinsic Gain

For the BJT, we can derive a formula for the intrinsic gain $A_0 = g_m r_o$ by using the following formulas for g_m and r_o:

$$g_m = \frac{I_C}{V_T} \tag{8.38}$$

$$r_o = \frac{V_A}{I_C} \tag{8.39}$$

The result is

$$A_0 = g_m r_o = \frac{V_A}{V_T} \tag{8.40}$$

Thus A_0 is simply the ratio of the Early voltage V_A, which is a technology-determined parameter, and the thermal voltage V_T, which is a physical parameter (approximately 0.025 V at room temperature). The value of V_A ranges from 5 V to 35 V for modern IC fabrication processes to 100 V to 130 V for the older, so-called high-voltage processes (see Appendix G). As a result, the value of A_0 will be in the range of 200 V/V to 5000 V/V, with the lower values characteristic of modern small-feature-size devices. It is important to note that for a given bipolar-transistor fabrication process, A_0 is independent of the transistor junction area and of its bias current. This is not the case for the MOSFET, as we shall now see.

Recall from our study of the MOSFET g_m in Section 7.2 that there are three possible expressions for g_m. Two of these are particularly useful for our purposes here:

$$g_m = \frac{I_D}{V_{OV}/2} \tag{8.41}$$

$$g_m = \sqrt{2\mu_n C_{ox}(W/L)}\sqrt{1_D} \tag{8.42}$$

For the MOSFET r_o we have

$$r_o = \frac{V_A}{I_D} = \frac{V'_A L}{I_D} \tag{8.43}$$

where V_A is the Early voltage and V'_A is the technology-dependent component of the Early voltage. Utilizing each of the g_m expressions together with the expression for r_o, we obtain for A_0,

$$A_0 = \frac{V_A}{V_{OV}/2} \tag{8.44}$$

which can be expressed in the alternate forms

$$A_0 = \frac{2V'_A L}{V_{OV}} \tag{8.45}$$

and

$$A_0 = \frac{V_A'\sqrt{2(\mu_n C_{ox})(WL)}}{\sqrt{I_D}} \tag{8.46}$$

The expression in Eq. (8.44) is the one most directly comparable to that of the BJT (Eq. 8.40). Here, however, we note the following:

1. The quantity in the denominator is $V_{OV}/2$, which is a design parameter. Although the value of V_{OV} that designers use for modern submicron technologies has been steadily decreasing, it is still about 0.15 V to 0.3 V. Thus $V_{OV}/2$ is 0.075 V to 0.15 V, which is 3 to 6 times higher than V_T. Furthermore, there are reasons for selecting higher values for V_{OV} (to be discussed in later chapters).
2. The numerator quantity is both process dependent (through V_A') and device dependent (through L), and its value has been steadily decreasing with the scaling down of the technology (see Appendix K).
3. From Eq. (8.45) we see that for a given technology (i.e., a given value of V_A') the intrinsic gain A_0 can be increased by using a longer MOSFET and operating it at a lower V_{OV}. As usual, however, there are design trade-offs. For instance, we will see in Chapter 10 that increasing L and lowering V_{OV} result, independently, in decreasing the amplifier bandwidth.

As a result, the intrinsic gain realized in a MOSFET fabricated in a modern short-channel technology is only 10 V/V to 40 V/V, an order of magnitude lower than that for a BJT.

The alternative expression for the MOSFET A_0 given in Eq. (8.46) reveals a very interesting fact: For a given process technology (V_A' and $\mu_n C_{ox}$) and a given device (W and L), the intrinsic gain is inversely proportional to $\sqrt{I_D}$. This is illustrated in Fig. 8.14, which shows a typical plot for A_0 versus the bias current I_D. The plot confirms that the gain increases as the bias current is lowered. The gain, however, levels off at very low currents. This is because the MOSFET enters the **subthreshold region** of operation (Section 5.1.9), where it becomes very much like a BJT with an exponential current–voltage characteristic. The intrinsic gain then becomes constant, just like that of a BJT. Note, however, that although higher gain is

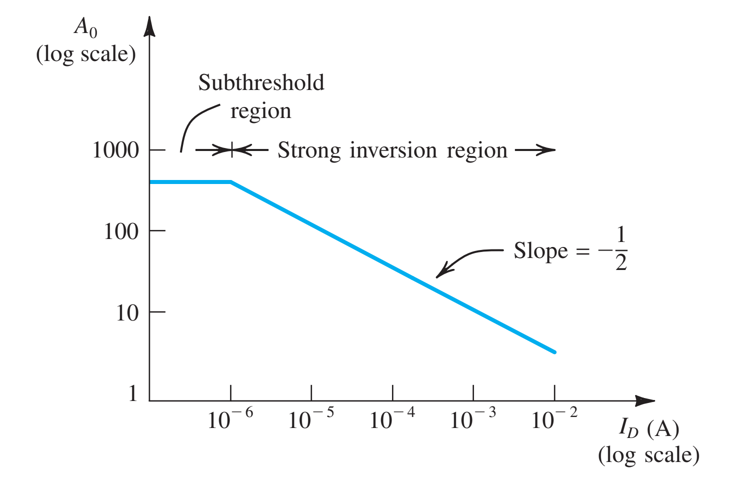


Figure 8.14 The intrinsic gain of the MOSFET versus bias current I_D. Outside the subthreshold region, this is a plot of $A_0 = V_A'\sqrt{2\mu_n C_{ox} WL/I_D}$ for the case: $\mu_n C_{ox} = 20\ \mu\text{A/V}^2$, $V_A' = 20$ V/μm, $L = 2$ μm, and $W = 20$ μm.

obtained at lower values of I_D, the price paid is a lower g_m (Eq. 8.42) and less ability to drive capacitive loads, and thus a decrease in bandwidth. This point will be studied in Chapter 10.

Example 8.2

We wish to compare the values of g_m, R_{in}, R_o, and A_0 for a CS amplifier that is designed using an NMOS transistor with $L = 0.4\ \mu m$ and $W = 4\ \mu m$ and fabricated in a 0.25-μm technology specified to have $\mu_n C_{ox} = 267\ \mu A/V^2$ and $V_A' = 10\ V/\mu m$, with those for a CE amplifier designed using a BJT fabricated in a process with $\beta = 100$ and $V_A = 10$ V. Assume that both devices are operating at a drain (collector) current of 100 μA.

Solution

For simplicity, we shall neglect the Early effect in the MOSFET in determining V_{OV}; thus,

$$I_D = \frac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)V_{OV}^2$$

$$100 = \frac{1}{2} \times 267 \times \left(\frac{4}{0.4}\right)V_{OV}^2$$

resulting in

$$V_{OV} = 0.27\ \text{V}$$

$$g_m = \frac{2I_D}{V_{OV}} = \frac{2 \times 0.1}{0.27} = 0.74\ \text{mA/V}$$

$$R_{in} = \infty$$

$$r_o = \frac{V_A' L}{I_D} = \frac{10 \times 0.4}{0.1} = 40\ \text{k}\Omega$$

$$R_o = r_o = 40\ \text{k}\Omega$$

$$A_0 = g_m r_o = 0.74 \times 40 = 29.6\ \text{V/V}$$

For the CE amplifier we have

$$g_m = \frac{I_C}{V_T} = \frac{0.1\ \text{mA}}{0.025\ \text{V}} = 4\ \text{mA/V}$$

$$R_{in} = r_\pi = \frac{\beta}{g_m} = \frac{100}{4} = 25\ \text{k}\Omega$$

$$r_o = \frac{V_A}{I_C} = \frac{10}{0.1} = 100\ \text{k}\Omega$$

$$R_o = r_o = 100\ \text{k}\Omega$$

$$A_0 = g_m r_o = 4 \times 100 = 400\ \text{V/V}$$

EXERCISE

8.7 A CS amplifier utilizes an NMOS transistor with $L = 0.36\ \mu$m and $W/L = 10$; it was fabricated in a 0.18-μm CMOS process for which $\mu_n C_{ox} = 387\ \mu\text{A/V}^2$ and $V_A' = 5$ V/μm. Find the values of g_m and A_0 obtained at $I_D = 10\ \mu$A, 100 μA, and 1 mA.
Ans. 0.28 mA/V, 50 V/V; 0.88 mA/V, 15.8 V/V; 2.78 mA/V, 5 V/V

8.3.3 Effect of the Output Resistance of the Current-Source Load

The current-source load of the CS amplifier in Fig. 8.13(a) can be implemented using a PMOS transistor biased in the saturation region to provide the required current I, as shown in Fig. 8.15(a). We can use the large-signal MOSFET model (Section 5.2, Fig. 5.18) to model

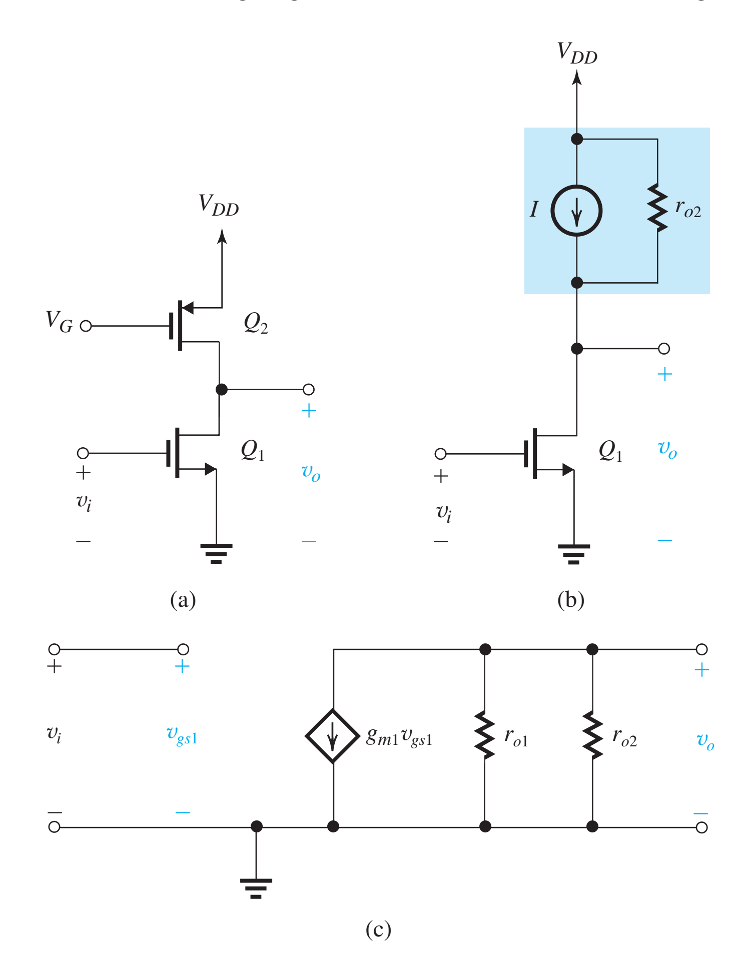


Figure 8.15 (**a**) The CS amplifier with the current-source load implemented with a p-channel MOSFET Q_2; (**b**) the circuit with Q_2 replaced with its large-signal model; and (**c**) small-signal equivalent circuit of the amplifier.

Q_2 as shown in Fig. 8.15(b), where

$$I = \frac{1}{2}(\mu_p C_{ox})\left(\frac{W}{L}\right)_2 \left[V_{DD} - V_G - |V_{tp}|\right]^2 \tag{8.47}$$

and

$$r_{o2} = \frac{|V_{A2}|}{I} \tag{8.48}$$

Thus the current-source load no longer has an infinite resistance; rather, it has a finite output resistance r_{o2}. This resistance will in effect appear in parallel with r_{o1}, as shown in the amplifier equivalent-circuit model in Fig. 8.15(c), from which we obtain

$$A_v \equiv \frac{v_o}{v_i} = -g_{m1}(r_{o1} \| r_{o2}) \tag{8.49}$$

Thus, not surprisingly, the finite output resistance of the current-source load reduces the magnitude of the voltage gain from $(g_{m1}r_{o1})$ to $g_{m1}(r_{o1} \| r_{o2})$. This reduction can be substantial. For instance, if Q_2 has an Early voltage equal to that of Q_1, $r_{o2} = r_{o1}$ and the gain is reduced by half,

$$A_v = -\frac{1}{2}g_m r_o \tag{8.50}$$

Finally, we note that a similar development can be used for the bipolar case.

Example 8.3

A practical circuit implementation of the common-source amplifier is shown in Fig. 8.16(a). Here the current-source transistor Q_2 is the output transistor of a current mirror formed by Q_2 and Q_3 and fed with a reference current I_{REF}. The NMOS version of this current source was studied in Section 8.1. Assume that Q_2 and Q_3 are matched. To be able to clearly see the region of v_I over which the circuit operates as an almost-linear amplifier, determine the voltage-transfer characteristic (VTC), that is, v_O versus v_I.

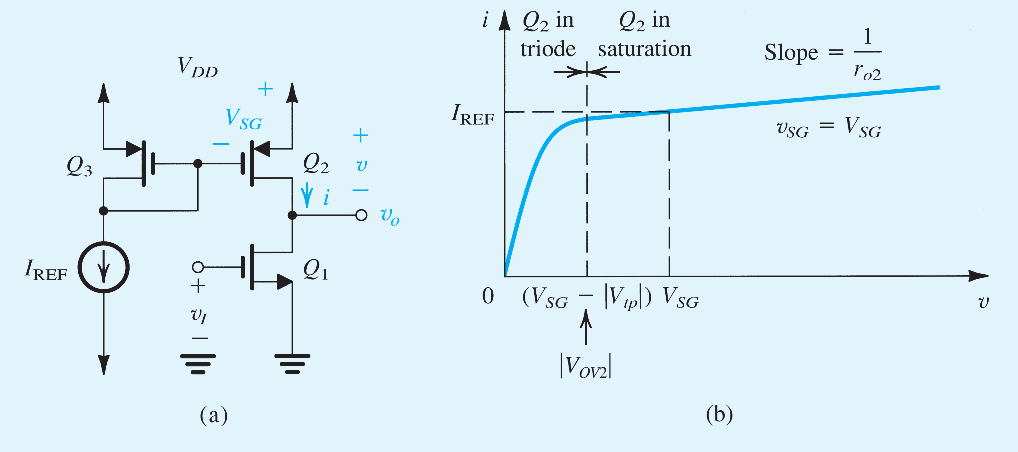


Figure 8.16 Practical implementation of the common-source amplifier: (**a**) circuit; (**b**) i–v characteristic of the active-load Q_2; (**c**) graphical construction to determine the transfer characteristic; (**d**) transfer characteristic.

Example 8.3 *continued*

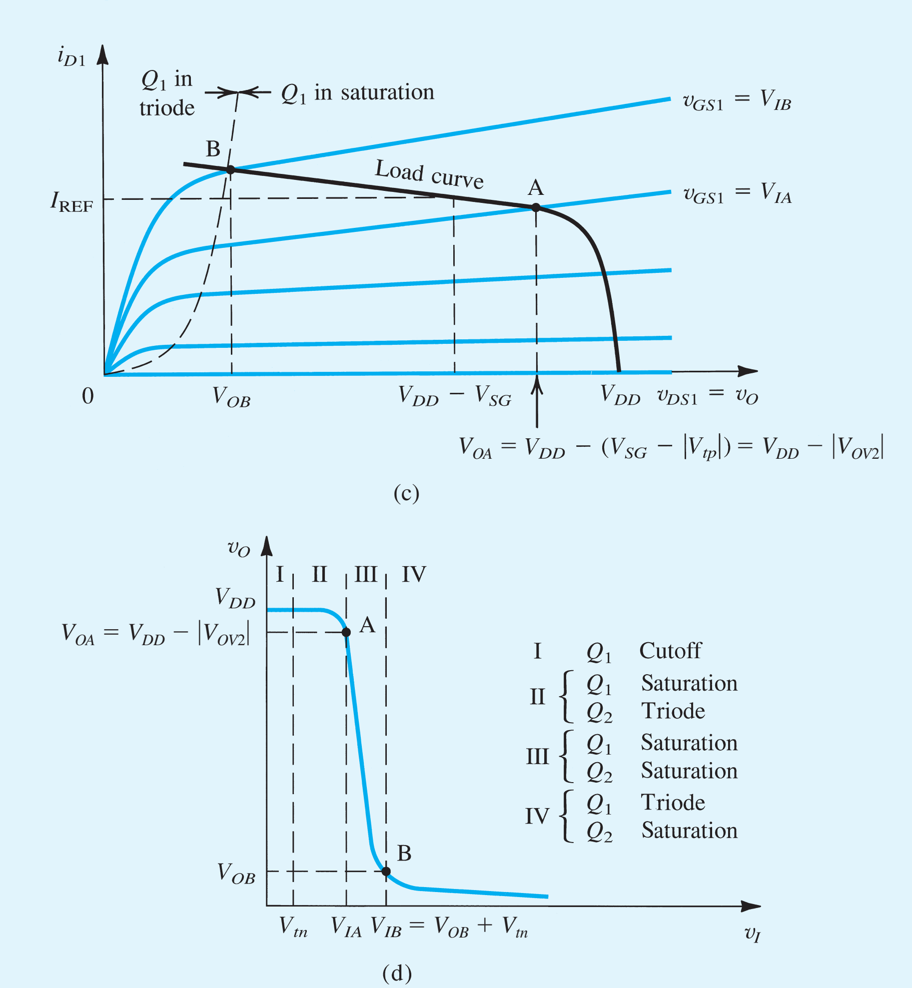


Figure 8.16 *continued*

Solution

First we concern ourselves with the current mirror, with the objective of determining the i–v characteristic of the current source Q_2. Toward that end, we note that the current I_{REF} flows through the diode-connected transistor Q_3 and thus determines V_{SG} of Q_3, which is in turn applied between the source and the gate of Q_2. Thus, the i–v characteristic of the current source Q_2 will be the i_D–v_{SD} characteristic curve of Q_2 obtained for $v_{SG} = V_{SG}$. This is shown in Fig. 8.16(b), where we note that i will be equal to I_{REF} at one point only, namely, at $v_{SD2} = V_{SG}$, this being the only point at which the two matched transistors Q_2 and Q_3 have identical operating conditions. We also observe the effect of channel-length modulation in Q_2 (the Early effect), which is modeled by the finite output resistance r_{o2}. Finally, note that Q_2 operates as

a current source when v is equal to or greater than $|V_{OV2}| = V_{SG} - |V_{tp}|$. This in turn is obtained when $v_O \leq V_{DD} - |V_{OV2}|$. This is the maximum permitted value of the output voltage v_O.

Now, with the i–v characteristic of the current-source load Q_2 in hand, we can proceed to determine v_O versus v_I. Figure 8.16(c) shows a graphical construction for doing this. It is based on the graphical analysis method employed in Section 7.1.6 except that here the load line is not a straight line but is the i–v characteristic curve of Q_2 shifted along the v_O axis by V_{DD} volts and "flipped around." The reason for this is that

$$v_O = V_{DD} - v$$

The term V_{DD} necessitates the shift, and the minus sign of v gives rise to the "flipping around" of the load curve.

The graphical construction of Fig. 8.16(c) can be used to determine v_O for every value of v_I, point by point: The value of v_I determines the particular characteristic curve of Q_1 on which the operating point lies. The operating point will be at the intersection of this particular graph and the load curve. The horizontal coordinate of the operating point then gives the value of v_O.

Proceeding in the manner just explained, we obtain the VTC shown in Fig. 8.16(d). As indicated, it has four distinct segments, labeled I, II, III, and IV. Each segment is obtained for one of the four combinations of the modes of operation of Q_1 and Q_2, which are also indicated in the diagram. Note that we have labeled two important break points on the transfer characteristic (A and B) in correspondence with the intersection points (A and B) in Fig. 8.16(c). We urge the reader to carefully study the transfer characteristic and its various details.

Not surprisingly, segment III is the one of interest for amplifier operation. Observe that in region III the transfer curve is almost linear and is very steep, indicating large voltage gain. In region III both the amplifying transistor Q_1 and the load transistor Q_2 are operating in saturation. The end points of region III are A and B: At A, defined by $v_O = V_{DD} - |V_{OV2}|$, Q_2 enters the triode region, and at B, defined by $v_O = v_I - V_{tn}$, Q_1 enters the triode region. When the amplifier is biased at a point in region III, the small-signal voltage gain can be determined as we have done in Fig. 8.15(c). The question remains as to how we are going to guarantee that the dc component of v_I will have such a value that will result in operation in region III. That is why overall negative feedback is needed, as will be demonstrated later.

Before leaving this example it is useful to reiterate that the upper limit of the amplifier region (i.e., point A) is defined by $V_{OA} = V_{DD} - |V_{OV2}|$ and the lower limit (i.e., point B) is defined by $V_{OB} = V_{OV1}$, where V_{OV1} can be approximately determined by assuming that $I_{D1} \simeq I_{REF}$. A more precise value for V_{OB} can be obtained by taking into account the Early effect in both Q_1 and Q_2, as will be demonstrated in the next example.

Example 8.4

Consider the CMOS common-source amplifier in Fig. 8.16(a) for the case $V_{DD} = 3$ V, $V_{tn} = |V_{tp}| = 0.6$ V, $\mu_n C_{ox} = 200\ \mu A/V^2$, and $\mu_p C_{ox} = 65\ \mu A/V^2$. For all transistors, $L = 0.4\ \mu m$ and $W = 4\ \mu m$. Also, $V_{An} = 20$ V, $|V_{Ap}| = 10$ V, and $I_{REF} = 100\ \mu A$. Find the small-signal voltage gain. Also, find the coordinates of the extremities of the amplifier region of the transfer characteristic— that is, points A and B.

Example 8.4 *continued*

Solution

$$g_{m1} = \sqrt{2k_n'\left(\frac{W}{L}\right)_1 I_{REF}}$$
$$= \sqrt{2 \times 200 \times \frac{4}{0.4} \times 100} = 0.63 \text{ mA/V}$$
$$r_{o1} = \frac{V_{An}}{I_{D1}} = \frac{20 \text{ V}}{0.1 \text{ mA}} = 200 \text{ k}\Omega$$
$$r_{o2} = \frac{|V_{Ap}|}{I_{D2}} = \frac{10 \text{ V}}{0.1 \text{ mA}} = 100 \text{ k}\Omega$$

Thus,

$$A_v = -g_{m1}(r_{o1} \parallel r_{o2})$$
$$= -0.63(\text{mA/V}) \times (200 \parallel 100)(\text{k}\Omega) = -42 \text{ V/V}$$

Approximate values for the extremities of the amplifier region of the transfer characteristic [region III in Fig. 8.16(d)] can be determined as follows: Neglecting the Early effect, all three transistors are carrying equal currents I_{REF}, and thus we can determine the overdrive voltages at which they are operating. Transistors Q_2 and Q_3 will have equal overdrive voltages, $|V_{OV3}|$, determined from

$$I_{D3} = I_{REF} \simeq \frac{1}{2}(\mu_p C_{ox})\left(\frac{W}{L}\right)_3 |V_{OV3}|^2$$

Substituting $I_{REF} = 100\ \mu\text{A}$, $\mu_p C_{ox} = 65\ \mu\text{A/V}^2$, $(W/L)_3 = 4/0.4 = 10$ results in

$$|V_{OV3}| = 0.55 \text{ V}$$

Thus,

$$V_{OA} = V_{DD} - |V_{OV3}| = 2.45 \text{ V}$$

Next we determine $|V_{OV1}|$ from

$$I_{D1} \simeq I_{REF} \simeq \frac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)_1 V_{OV1}^2$$

Substituting $I_{REF} = 100\ \mu\text{A}, \mu_n C_{ox} = 200\ \mu\text{A/V}^2, (W/L)_1 = 4/0.4 = 10$ results in

$$V_{OV1} = 0.32 \text{ V}$$

Thus,

$$V_{OB} = V_{OV1} = 0.32 \text{ V}.$$

More precise values for V_{OA} and V_{OB} can be determined by taking the Early effect in all transistors into account as follows.

First, we determine V_{SG} of Q_2 and Q_3 corresponding to $I_{D3} = I_{\text{REF}} = 100\ \mu\text{A}$ using

$$I_{D3} = \frac{1}{2}k_p'\left(\frac{W}{L}\right)_3 \left(V_{SG} - |V_{tp}|\right)^2\left(1 + \frac{V_{SD}}{|V_{Ap}|}\right)$$

Thus,

$$100 = \frac{1}{2} \times 65\left(\frac{4}{0.4}\right)|V_{OV3}|^2\left(1 + \frac{0.6 + |V_{OV3}|}{10}\right) \tag{8.51}$$

where $|V_{OV3}|$ is the magnitude of the overdrive voltage at which Q_3 and Q_2 are operating, and we have used the fact that, for Q_3, $V_{SD} = V_{SG}$. Equation (8.51) can be manipulated to the form

$$0.29 = |V_{OV3}|^2\left(1 + 0.09|V_{OV3}|\right)$$

which by a trial-and-error process yields

$$|V_{OV3}| = 0.526\ \text{V}$$

Thus,

$$V_{SG} = 0.6 + 0.526 = 1.126\ \text{V}$$

and

$$V_{OA} = V_{DD} - V_{OV3} = 2.47\ \text{V}$$

To find the corresponding value of v_I, V_{IA}, we derive an expression for v_O versus v_I in region III. Noting that in region III, Q_1 and Q_2 are in saturation and obviously conduct equal currents, we can write

$$i_{D1} = i_{D2}$$

$$\frac{1}{2}k_n'\left(\frac{W}{L}\right)_1 \left(v_I - V_{tn}\right)^2\left(1 + \frac{v_O}{|V_{An}|}\right) = \frac{1}{2}k_p'\left(\frac{W}{L}\right)_2 \left(V_{SG} - |V_{tp}|\right)^2\left(1 + \frac{V_{DD} - v_O}{|V_{Ap}|}\right)$$

Substituting numerical values, we obtain

$$8.55\left(v_I - 0.6\right)^2 = \frac{1 - 0.08v_O}{1 + 0.05v_O} \tag{8.52}$$

This is the equation of segment III of the transfer characteristic. Although it includes v_I^2, the reader should not be alarmed: Because region III is very narrow, v_I changes very little, and the characteristic is nearly linear. Substituting $v_O = 2.47$ V gives the corresponding value of v_I; that is, $V_{IA} = 0.89$ V. To determine the coordinates of B, we note that they are related by $V_{OB} = V_{IB} - V_{tn}$. Substituting in Eq. (8.52) and solving by trial and error gives $V_{IB} = 0.935$ V and $V_{OB} = 0.335$ V. The width of the amplifier region is therefore

$$\Delta v_I = V_{IB} - V_{IA} = 0.045\ \text{V}$$

and the corresponding output range is

$$\Delta v_O = V_{OB} - V_{OA} = -2.135\ \text{V}$$

Example 8.4 *continued*

Thus, the "large-signal" voltage gain is

$$\frac{\Delta v_O}{\Delta v_I} = -\frac{2.135}{0.045} = -47.4 \text{ V/V}$$

which is reasonably close to the small-signal value of –42, indicating that segment III of the transfer characteristic is quite linear.

EXERCISES

8.8 A CMOS common-source amplifier such as that in Fig. 8.16(a), fabricated in a 0.18-μm technology, has $W/L = 7.2\ \mu\text{m}/0.36\ \mu\text{m}$ for all transistors, $k_n' = 387\ \mu\text{A/V}^2$, $k_p' = 86\ \mu\text{A/V}^2$, $I_{REF} = 100\ \mu\text{A}$, $V_{An}' = 5\text{ V}/\mu\text{m}$, and $|V_{Ap}'| = 6\text{ V}/\mu\text{m}$. Find g_{m1}, r_{o1}, r_{o2}, and the voltage gain.
Ans. 1.24 mA/V; 18 kΩ; 21.6 kΩ; –12.2 V/V

8.9 Consider the active-loaded CE amplifier when the constant-current source I is implemented with a *pnp* transistor. Let $I = 0.1$ mA, $|V_A| = 50$ V (for both the *npn* and the *pnp* transistors), and $\beta = 100$. Find R_{in}, r_o (for each transistor), g_m, A_0, and the amplifier voltage gain.
Ans. 25 kΩ; 0.5 MΩ; 4 mA/V; 2000 V/V; –1000 V/V

8.3.4 Increasing the Gain of the Basic Cell

We conclude this section by considering a question: How can we increase the voltage gain obtained from the basic gain cell? The answer lies in finding a way to raise the level of the output resistance of both the amplifying transistor and the load transistor. That is, we seek a circuit that passes the current $g_m v_i$ provided by the amplifying transistor right through, but increases the resistance from r_o to a much larger value. This requirement is illustrated in Fig. 8.17. Figure 8.17(a) shows the CS amplifying transistor Q_1 together with its output equivalent circuit. Note that for the time being we are not showing the load device. In Fig. 8.17(b) we have inserted a shaded box between the drain of Q_1 and a new output terminal labeled d_2. Here again we are not showing the load to which d_2 will be connected. Our "black box" takes in the output current of Q_1 and passes it to the output; thus at its output we have the equivalent circuit shown, consisting of the same controlled source $g_{m1} v_i$ but with the output resistance increased by a factor K.

Now, what does the black box really do? Since it passes the *current* but *raises* the resistance level, it is a **current buffer**. It is the dual of the voltage buffer (the source and emitter followers), which passes the *voltage* but *lowers* the resistance level.

Now searching our repertoire of transistor amplifier configurations studied in Section 7.3, the only candidate for implementing this current-buffering action is the common-gate (or common-base in bipolar) amplifier. Indeed, recall that the CG and CB circuits have a unity current gain. What we have not yet investigated, however, is their resistance transformation property. We shall do this in the next section.

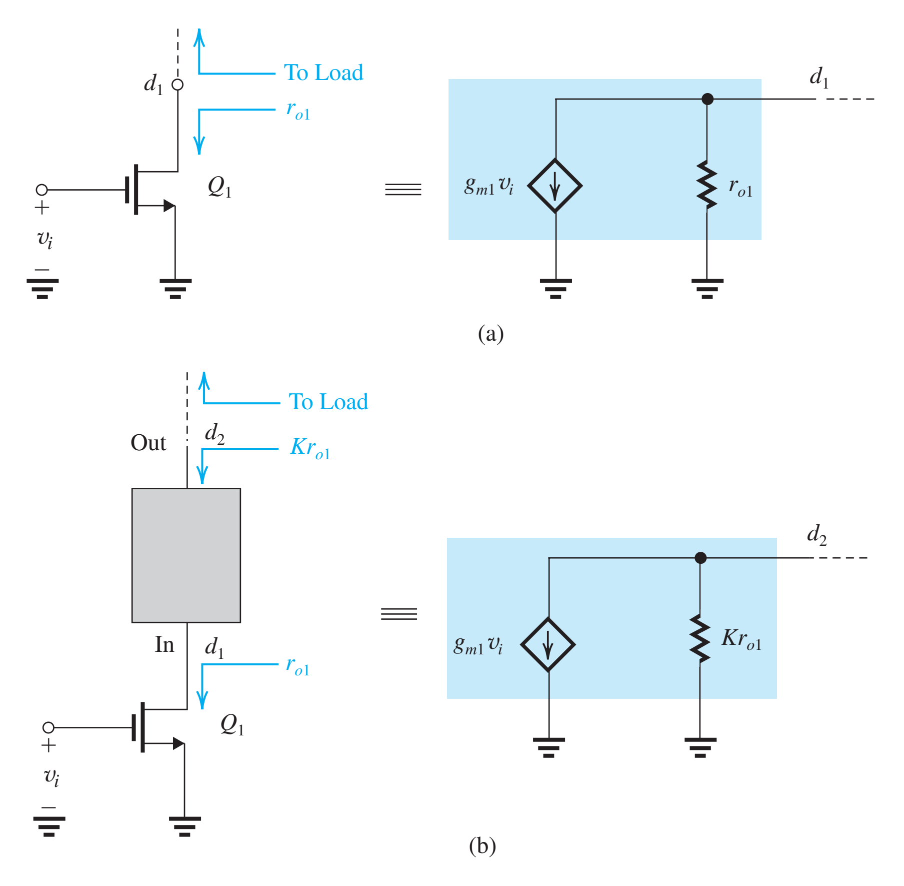


Figure 8.17 To increase the voltage gain realized in the basic gain cell shown in (**a**), a functional block, shown as a black box in (**b**), is connected between d_1 and the load. This new block is required to pass the current $g_{m1}v_i$ right through but raise the resistance level by a factor K. The functional block is a current buffer and can be realized with a common-gate transistor, as demonstrated in the next section.

8.4 The Common-Gate and Common-Base Amplifiers

In this section we study the IC versions of the CG and CB amplifier configurations. This study differs in a significant way from that of the discrete-circuit versions (Section 7.3.5) because here we have to take into account the output resistance of the transistor, r_o. In the following, we show that both the CG and CB configurations provide excellent implementations of the current buffer discussed in the previous section.

8.4.1 The CG Circuit

Figure 8.18(a) shows a CG amplifier with the biasing arrangement shown only partially. The amplifier is fed with a signal source v_{sig} having a resistance R_s, and it has a load resistance R_L. The latter is usually implemented using a PMOS current source, as discussed earlier.

To characterize the signal performance of the CG amplifier, we show in Fig. 8.18(b) the circuit with the dc voltages eliminated. Observe that because the gate current is zero, the input

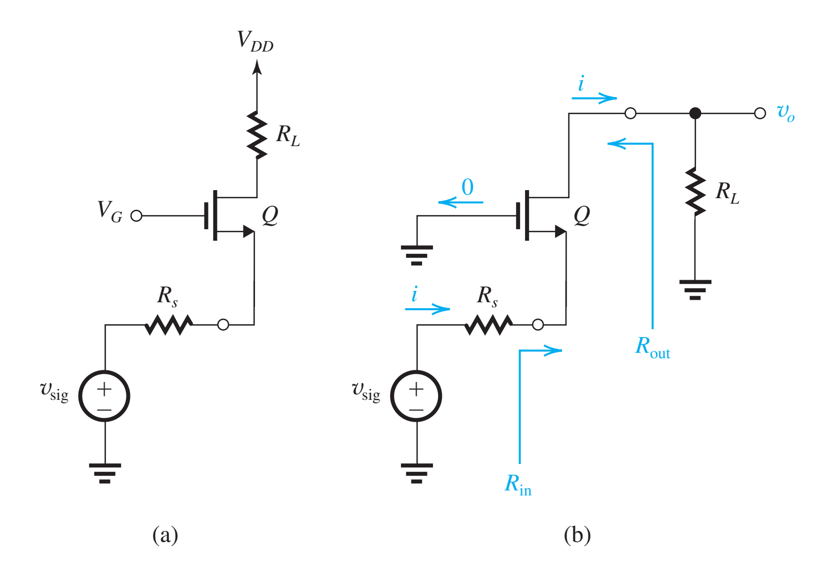


Figure 8.18 (**a**) A CG amplifier with the bias arrangement only partially shown. (**b**) The circuit with the dc sources eliminated.

current i passes through to the drain and on to the load—the first requirement of a current buffer.

Input Resistance The input resistance R_{in} can be found using the circuit of Fig. 8.19. Here we have employed the T model of the MOSFET and applied a test voltage v_x to the input. The input resistance is given by

$$R_{in} \equiv \frac{v_x}{i_x}$$

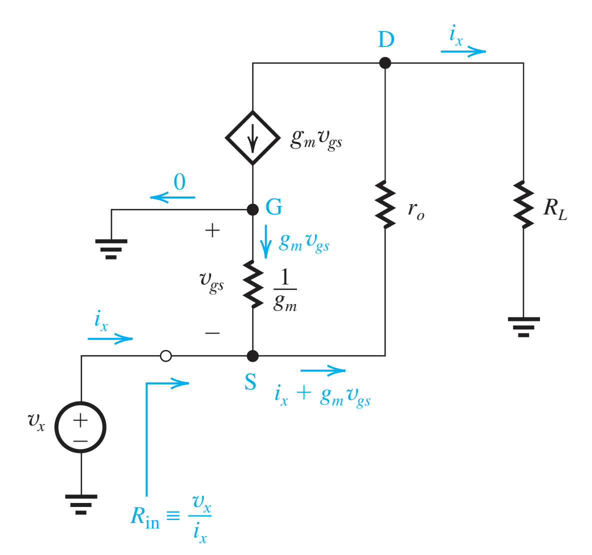


Figure 8.19 Determining the input resistance R_{in} of the CG amplifier.

The analysis proceeds as follows.

A node equation at the input yields the current in r_o as $(i_x + g_m v_{gs})$. A node equation at the output shows that the current through R_L is i_x. Next, a loop equation for the loop comprising v_x, r_o, and R_L gives

$$v_x = (i_x + g_m v_{gs})r_o + i_x R_L$$

Since the voltage at the source node v_x is equal to $-v_{gs}$, we can replace v_{gs} by $-v_x$ and rearrange terms to obtain $R_{\text{in}} \equiv v_x/i_x$,

$$R_{\text{in}} = \frac{r_o + R_L}{1 + g_m r_o} \tag{8.53}$$

For $g_m r_o \gg 1$,

$$R_{\text{in}} \simeq \frac{1}{g_m} + \frac{R_L}{g_m r_o} \tag{8.54}$$

This is a very interesting result. First, it shows that if r_o is infinite, as was the case in our analysis of the discrete CG amplifier in Section 7.3.5, then R_{in} reduces to $1/g_m$, verifying the result we found there. If r_o cannot be neglected, as is always the case in IC amplifiers, we see that the input resistance depends on R_L in an interesting fashion: The load resistance R_L is transformed to the input by dividing it by the intrinsic gain $A_0 = g_m r_o$. Thus, even as R_L is increased, this impedance transformation property ensures that R_{in} remains relatively low, an important characteristic of a current buffer.

Output Resistance To obtain the output resistance R_{out} we utilize the circuit shown in Fig. 8.20. Here we have short circuited v_{sig} but left the source resistance R_s, and applied a test voltage v_x to the output. The output resistance is given by

$$R_{\text{out}} = \frac{v_x}{i_x}$$

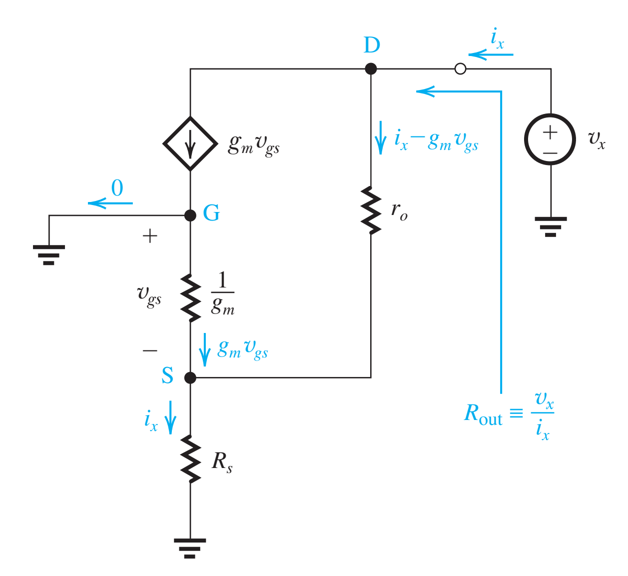


Figure 8.20 Determining the output resistance R_{out} of the CG amplifier.

The analysis proceeds as follows.

A node equation at the drain gives the current through r_o as $(i_x - g_m v_{gs})$. A node equation at the source gives the current in R_s as i_x. Next, a loop equation for the loop comprising v_x, r_o, and R_s gives

$$v_x = (i_x - g_m v_{gs})r_o + i_x R_s \tag{8.55}$$

Finally, we observe that the voltage at the source terminal is $-v_{gs}$ and can also be expressed as $i_x R_s$, thus

$$v_{gs} = -i_x R_s$$

Substituting this value for v_{gs} into Eq. (8.55) and rearranging terms to obtain $R_{\text{out}} \equiv v_x / i_x$ yields

$$R_{\text{out}} = r_o + R_s + g_m r_o R_s \tag{8.56}$$

which can be written in the alternate form

$$R_{\text{out}} = r_o + (1 + g_m r_o)R_s \tag{8.57}$$

For $g_m r_o \gg 1$,

$$R_{\text{out}} \simeq r_o + (g_m r_o)R_s \tag{8.58}$$

and if we also have $g_m R_s \gg$ then

$$R_{\text{out}} \simeq (g_m r_o)R_s \tag{8.59}$$

Equation (8.58) indicates that the output resistance of the CG amplifier includes, in addition to the transistor's r_o, a component related to the resistance in the source load R_s. The significant point is that the CG amplifier transforms the source resistance R_s to the output by multiplying it by the intrinsic gain $A_0 = g_m r_o$. This impedance transformation is the inverse to that observed from output to input. Now, if R_s is large then the output resistance of the CG circuit can be very large; this also is an important characteristic of a current buffer.

To summarize: the CG circuit has a unity current gain; a low input resistance, obtained by dividing R_L by $g_m r_o$; and a high output resistance, obtained by multiplying R_s by $g_m r_o$. Thus it makes for an excellent current buffer and can be used to implement the shaded functional box in Fig. 8.17. As a useful summary, Fig. 8.21 illustrates the impedance transformation properties of the common-gate amplifier.

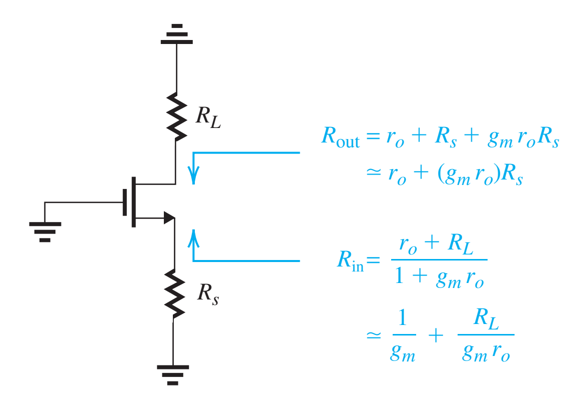


Figure 8.21 The impedance transformation properties of the common-gate amplifier. Depending on the values of R_s and R_L, we can sometimes write $R_{\text{in}} \simeq R_L/(g_m r_o)$ and $R_o \simeq (g_m r_o)R_s$. However, such approximations are not always justified.

EXERCISES

8.10 For the CG amplifier in Fig. 8.18, show that the voltage gain is given by

$$\frac{v_o}{v_{sig}} = \frac{R_L}{R_s + R_{in}}$$

8.11 For a CG amplifier for which $g_m r_o \gg 1$, find R_{in} for the following cases: $R_L = 0$; r_o; $(g_m r_o)r_o$; ∞.
Ans. $\frac{1}{g_m}$; $\frac{2}{g_m}$; r_o; ∞

8.12 For a CG amplifier for which $g_m r_o \gg 1$, find R_{out} for the following cases: $R_s = 0$; r_o; $(g_m r_o)r_o$; ∞.
Ans. r_o; $(g_m r_o)r_o$; $(g_m r_o)^2 r_o$; ∞

8.4.2 Output Resistance of a CS Amplifier with a Source Resistance

In Section 7.3.4 we discussed some of the benefits that are obtained when a resistance R_s is included in the source lead of a CS amplifier, as in Fig. 8.22. Such a resistance is referred to as a source-degeneration resistance because of its action in reducing the effective transconductance of the CS stage to $g_m/(1 + g_m R_s)$, that is, by a factor $(1 + g_m R_s)$. This also is the factor by which a number of performance parameters are increased, such as linearity and bandwidth (as will be seen in Chapter 10). At this juncture we simply wish to point out that the expression we derived above for the output resistance of the CG amplifier applies directly to the case of a source-degenerated CS amplifier. This is because when we determine R_o, we ground the input terminal, making transistor Q appear as a CG transistor. Thus R_o is given by Eq. (8.56), namely,

$$R_o = r_o + R_s + g_m r_o R_s \tag{8.60}$$

Since $g_m r_o \gg 1$, the second term on the right-hand side will be much lower than the third and can be neglected, resulting in

$$R_o \simeq (1 + g_m R_s) r_o \tag{8.61}$$

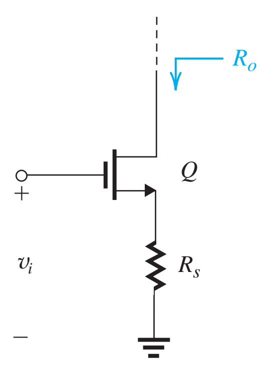


Figure 8.22 The output resistance expression of the CG amplifier can be used to find the output resistance of a source-degenerated common-source amplifier. Here, a useful interpretation of the result is that R_s increases the output resistance by the factor $(1 + g_m R_s)$.

Thus source degeneration increases the output resistance of the CS amplifier from r_o to $(1+g_mR_s)r_o$, again by the same factor $(1+g_mR_s)$. In Chapter 11, we will find that R_s introduces negative (degenerative) feedback of an amount $(1+g_mR_s)$.

EXERCISE

8.13 Given that source degeneration reduces the transconductance of a CS amplifier from g_m to approximately $g_m/(1+g_mR_s)$ and increases its output resistance by approximately the same factor, what happens to the open-circuit voltage gain A_{vo}? Now, find an expression for A_v when a load resistance R_L is connected to the output.

Ans. A_{vo} remains constant at g_mr_o:

$$A_v = (g_mr_o)\frac{R_L}{R_L+(1+g_mR_s)r_o} \tag{E.8.13}$$

8.4.3 The Body Effect

Since in the CG amplifier the source cannot be connected to the substrate, the body effect (see Section 5.4) plays a role in the operation of the CG amplifier. It turns out, however, that taking the body effect into account in the analysis of the CG circuit is a very simple matter. To see how this can be done, refer to Fig. 8.23(a) and recall that the body terminal acts as another gate for the MOSFET. Thus, just as a signal voltage v_{gs} between the gate and the source gives

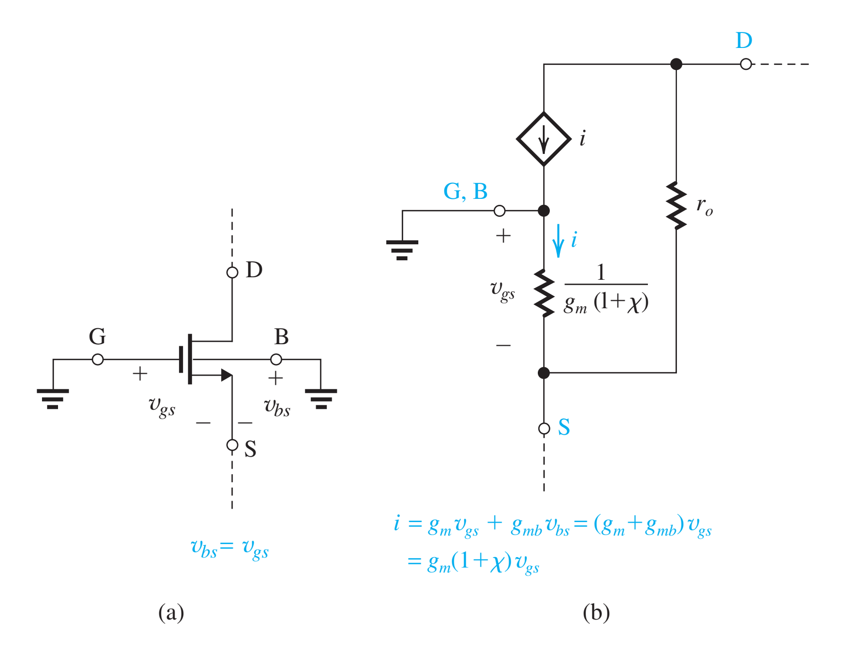


Figure 8.23 The body effect can be easily taken into account in the analysis of the CG circuit by replacing g_m by $(1+\chi)g_m$, where $\chi = g_{mb}/g_m = 0.1$ to 0.2.

rise to a drain current signal $g_m v_{gs}$, a signal voltage v_{bs} between the body and the source gives rise to a drain current signal $g_{mb} v_{bs}$. Thus the drain signal current becomes $(g_m v_{gs} + g_{mb} v_{bs})$, where the body transconductance g_{mb} is a small fraction χ of g_m; $g_{mb} = \chi g_m$ and $\chi = 0.1$ to 0.2. For the CG circuit, $v_{bs} = v_{gs}$, thus the two current signals can be combined as $(g_m + g_{mb})v_{gs}$ or $g_m(1+\chi)v_{gs}$. Thus, the body effect can be taken into account by simply replacing g_m by $g_m(1+\chi)$ as illustrated in the T equivalent model shown in Fig. 8.23(b). Normally, however, we will not bother with the factor $(1+\chi)$ in our calculations.

8.4.4 The CB Circuit

Analysis of the CB amplifier parallels that of the CG amplifier except it is a little more involved because of the finite base current. Figure 8.24(a) shows a CB amplifier with the bias details only partially shown and with a load resistance R_L that is normally implemented with a *pnp* current source. The circuit, prepared for small-signal analysis, is shown in Fig. 8.24(b). Note that since $\alpha \simeq 1$ the current gain is nearly unity, an important characteristic of a current buffer.

Input Resistance The circuit for determining the input resistance R_{in} is shown in Fig. 8.25, which also shows the currents in all branches, obtained by writing node equations for the three nodes. Of special note is the use of the identity $g_m + \frac{1}{r_\pi} = \frac{1}{r_e}$ to obtain the current in the base as v_π / r_π. Writing a loop equation for the loop comprising v_x, r_o, and R_L, and replacing v_π by $-v_x$, results in the following expression for $R_{\text{in}} \equiv v_x / i_x$,

$$R_{\text{in}} = \frac{r_o + R_L}{1 + \dfrac{r_o}{r_e} + \dfrac{R_L}{(\beta+1)r_e}} \tag{8.62}$$

where we have utilized the relationship $r_\pi = (\beta+1)r_e$. Since $r_o \gg r_e$,

$$R_{\text{in}} \simeq r_e \frac{r_o + R_L}{r_o + \dfrac{R_L}{\beta+1}} \tag{8.63}$$

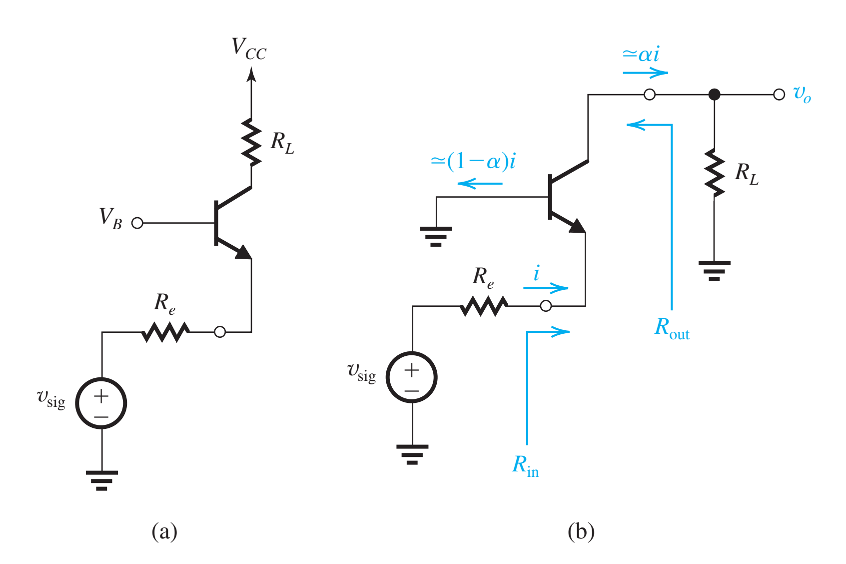


Figure 8.24 (**a**) A CB amplifier with the bias arrangement only partially shown. (**b**) The circuit with the dc sources eliminated.

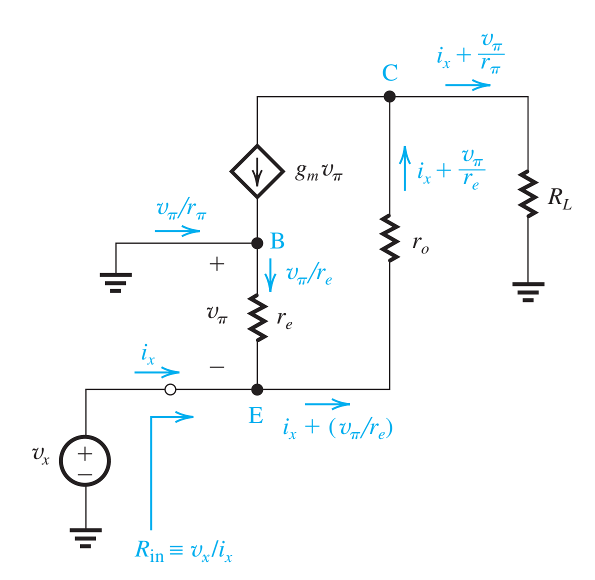


Figure 8.25 Determining the input resistance R_{in} of the CB amplifier.

Note that setting $r_o = \infty$ yields $R_{in} = r_e$, which is consistent with the case of the discrete-circuit CB amplifier studied in Section 7.3.5. Also, for $R_L = 0$, $R_{in} = r_e$. The value of R_{in} increases as R_L is raised, reaching a maximum of

$$R_{in}\big|_{max} = (\beta + 1)r_e = r_\pi, \text{ for } R_L = \infty \tag{8.64}$$

that is, for the amplifier operating open circuited. Finally, for $\frac{R_L}{\beta + 1} \ll r_o$, Eq. (8.63) can be approximated by

$$R_{in} \simeq r_e + \frac{R_L}{g_m r_o} \tag{8.65}$$

which is very similar to the case of the MOSFET (Eq. 8.54). We conclude by noting that the impedance transformation property of the CB circuit ensures that its input resistance is kept small, an important characteristic of a current buffer.

Output Resistance The determination of the output resistance R_{out} of the CB amplifier is illustrated in Fig. 8.26. The result is

$$R_{out} = r_o + (R_e \| r_\pi) + (R_e \| r_\pi)g_m r_o \tag{8.66}$$

which is very similar to the corresponding expression for the MOSFET case (Eq. 8.56) except that R_s is replaced by $(R_e \| r_\pi)$. The expression in Eq. (8.66) can be written in the alternate form

$$R_{out} = r_o + (1 + g_m r_o)(R_e \| r_\pi) \tag{8.67}$$

For $g_m r_o \gg 1$,

$$R_{out} \simeq r_o + (g_m r_o)(R_e \| r_\pi) \tag{8.68}$$

Thus, similar to the CG amplifier, the CB amplifier exhibits an impedance transformation property that raises the output resistance. Unlike the CG case, however, the output resistance of the CB circuit has an absolute maximum value obtained by setting $R_e = \infty$ as

$$R_{out}\big|_{max} = r_o + g_m r_o r_\pi = (\beta + 1)r_o \tag{8.69}$$

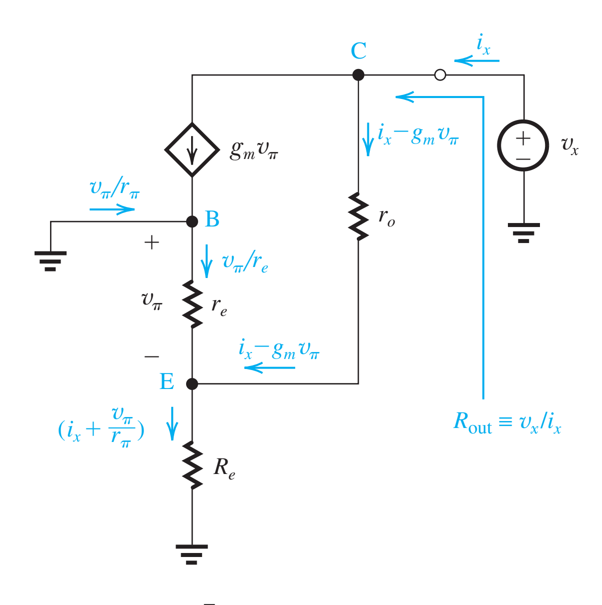


Figure 8.26 Determining the output resistance R_{out} of the CB amplifier.

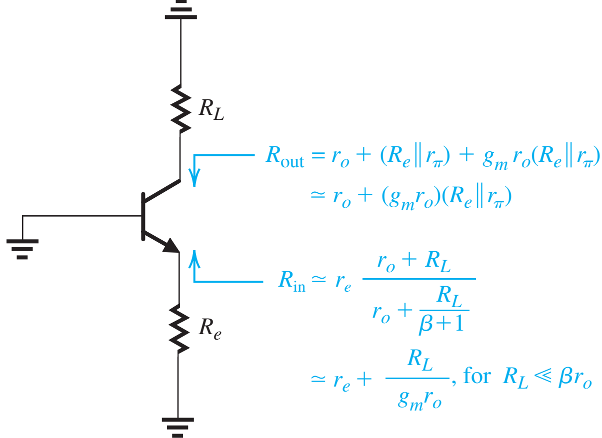


Figure 8.27 The impedance transformation properties of the CB amplifier. Note that for $\beta = \infty$, these formulas reduce to those for the MOSFET case (Fig. 8.21).

We conclude that the CB circuit has a current gain of nearly unity, a low input resistance, and a high output resistance; thus it makes for an excellent current buffer. The impedance transformation properties of the CB circuit are summarized in Fig. 8.27.

EXERCISES

8.14 For a CB amplifier, find approximate values for R_{in} for the following cases: $R_L = 0$; r_o; $(\beta + 1)r_o$; ∞.
Ans. r_e; $2r_e$; $\frac{1}{2}r_\pi$; r_π

8.15 For a CB amplifier, find approximate values for R_{out} for the following cases: $R_e = 0$; r_e; r_π; r_o; ∞.
Ans. r_o; $2r_o$; $(\frac{1}{2}\beta + 1)r_o$; $(\beta + 1)r_o$; $(\beta + 1)r_o$

8.4.5 Output Resistance of an Emitter-Degenerated CE Amplifier

As we have done in the MOS case, we shall adapt the expression for R_o derived for the CB amplifier (Eq. 8.68) for the case of a CE amplifier with a resistance R_e connected in its emitter, as shown in Fig. 8.28(a),

$$R_o \simeq r_o + g_m r_o (R_e \| r_\pi)$$

which can be written in the alternate form

$$R_o = [1 + g_m (R_e \| r_\pi)] r_o \tag{8.70}$$

Thus, emitter degeneration multiplies the transistor output resistance r_o by the factor $[1 + g_m (R_e \| r_\pi)]$. Note that this factor has a maximum value of $(1 + g_m r_\pi)$ or $(\beta + 1)$, obtained when $R_e \gg r_\pi$. Thus the theoretical maximum output resistance realized is $(\beta + 1) r_o$ and is achieved when the emitter is open circuited.

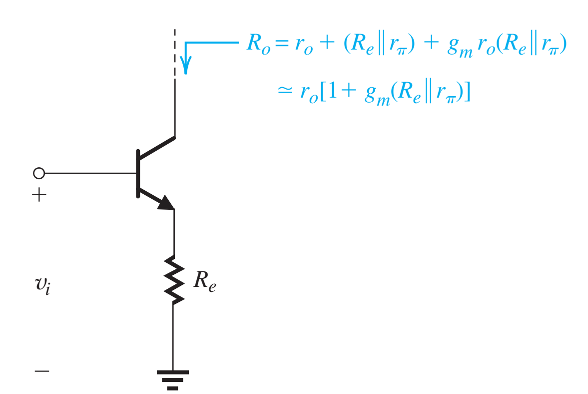


Figure 8.28 Output resistance of a CE amplifier with an emitter resistance R_e.

EXERCISE

8.16 Find the output resistance of a CE amplifier biased at $I_C = 1$ mA and having a resistance of 500 Ω connected in its emitter. Let $\beta = 100$ and $V_A = 10$ V. What is the value of the output resistance without degeneration?
Ans. 177 kΩ; 10 kΩ

8.5 The Cascode Amplifier

8.5.1 Cascoding

Cascoding refers to the use of a transistor connected in the common-gate (or the common-base) configuration to provide **current buffering** for the output of a common-source (or a common-emitter) amplifying transistor. Figure 8.29 illustrates the technique for the MOS case. Here the CS transistor Q_1 is the amplifying transistor and Q_2, connected in the CG

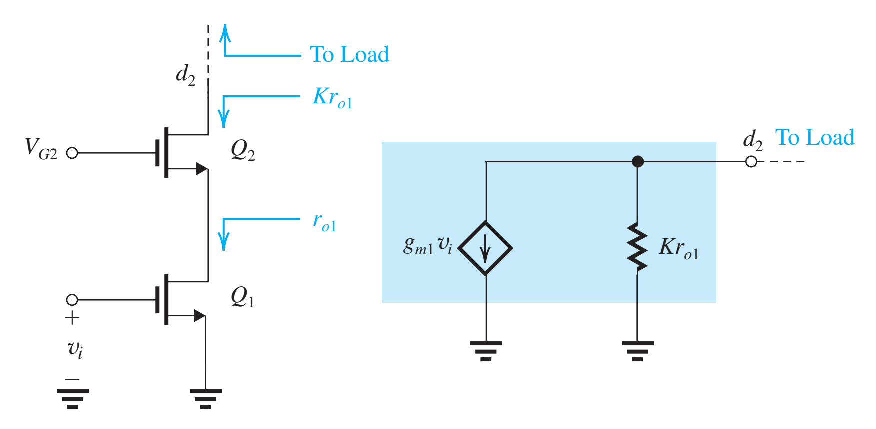


Figure 8.29 The current-buffering action of Fig. 8.17(b) is implemented using a transistor Q_2 connected in the CG configuration. Here V_{G2} is a dc bias voltage. The output equivalent circuit indicates that the CG transistor passes the current $g_{m1}v_i$ through but raises the resistance level by a factor K. Transistor Q_2 is called a cascode transistor.

configuration with a dc bias voltage V_{G2} (signal ground) at its gate, is the cascode transistor.[2] A similar arrangement applies for the bipolar case and will be considered later.

From our study of the CG amplifier characteristics, we can see that the cascode transistor passes the current $g_{m1}v_i$ to the output node while multiplying the resistance in its source (r_{o1} of Q_1) by a factor K. The result is the equivalent circuit of Fig. 8.29, which can be utilized to determine the voltage gain of the cascode amplifier for various load resistances. We shall consider the MOS cascode amplifier in detail next.

EXERCISE

8.17 Give an approximate value of the factor K of the circuit in Fig. 8.29.
Ans. $K \simeq g_{m2}r_{o2}$

8.5.2 The MOS Cascode Amplifier

The Ideal Case Figure 8.30(a) shows a MOS cascode amplifier loaded with an ideal constant-current source. The voltage gain realized can be found from the equivalent circuit in Fig. 8.30(b). Since the load is an ideal constant-current source, the load resistance is infinite. That is, the amplifier is operating with an open-circuit load, and the gain is

$$A_{vo} \equiv \frac{v_o}{v_i} = -g_{m1}R_o \tag{8.71}$$

[2] The name *cascode* is a carryover from the days of vacuum tubes and is a shortened version of "*casc*aded cath*ode*"; in the tube version, the anode of the amplifying tube (corresponding to the drain of Q_1) feeds the cathode of the cascode tube (corresponding to the source of Q_2).

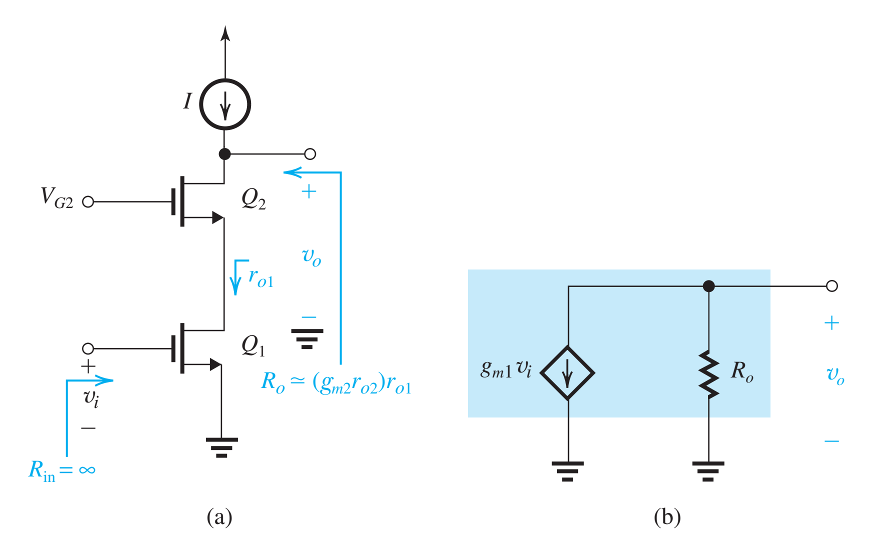


Figure 8.30 (**a**) A MOS cascode amplifier with an ideal current-source load; (**b**) equivalent circuit representation of the cascode output.

Now, since R_s of Q_2 is r_{o1}, the output resistance R_o is given by the approximate expression

$$R_o \simeq (g_{m2}r_{o2})r_{o1} \tag{8.72}$$

Substituting in Eq. (8.71) results in

$$A_{vo} = -(g_{m1}r_{o1})(g_{m2}r_{o2}) \tag{8.73}$$

For the case $g_{m1} = g_{m2} = g_m$ and $r_{o1} = r_{o2} = r_o$,

$$A_{vo} = -(g_m r_o)^2$$

$$= -A_0^2 \tag{8.74}$$

Thus cascoding increases the gain magnitude from A_0 to A_0^2.

Implementation of the Constant-Current Source Load If the current source load is implemented with a PMOS transistor (which can be part of a PMOS current mirror) as shown in Fig. 8.31(a), the load resistance R_L will be equal to the output resistance of Q_3, r_{o3},

$$R_L = r_{o3}$$

and the voltage gain of the cascode amplifier will be

$$A_v = -g_{m1}(R_o \| R_L)$$

$$= -g_{m1}(g_{m2}r_{o2}r_{o1} \| r_{o3}) \tag{8.75}$$

from which we can readily see that since $R_L \ll R_o$, the total resistance will be approximately equal to r_{o3} and the gain will be

$$A_v \simeq -g_{m1}r_{o3} \tag{8.76}$$

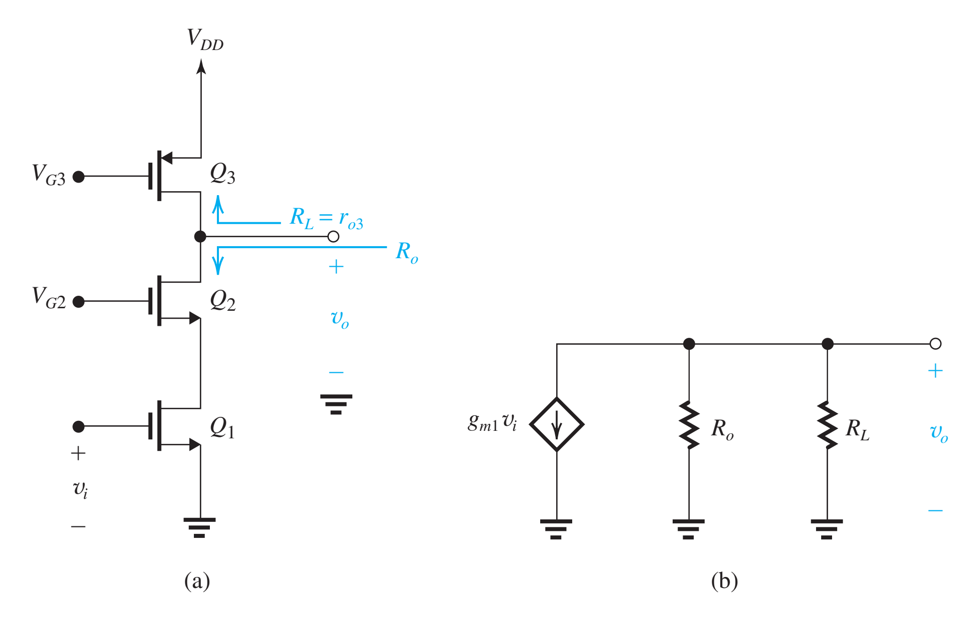


Figure 8.31 (**a**) A MOS cascode amplifier loaded in a simple PMOS current source Q_3. (**b**) Equivalent circuit at the amplifier output.

Thus the gain magnitude will be back to A_0, of the same order as that realized by a CS amplifier. In other words, the use of a simple current-source load with a relatively low output resistance has in effect destroyed the cascoding advantage of increased output resistance. Nevertheless, it turns out that this cascode amplifier, whose gain is of the same order as that of a CS amplifier, does in fact have a major advantage over the CS circuit: It exhibits a much wider bandwidth. We will demonstrate this point in Chapter 10.

The Use of a Cascode Current Source To realize a gain of the order of A_0^2, the load resistance R_L must be of the same order as R_o of the cascode amplifier. This can be achieved by using a cascode current source such as that shown in Fig. 8.32. Here Q_4 is the current-source transistor, and Q_3 is the CG cascode transistor. Voltages V_{G3} and V_{G4} are dc bias voltages. The cascode transistor Q_3 multiplies the output resistance of Q_4, r_{o4} by $(g_{m3}r_{o3})$ to provide an

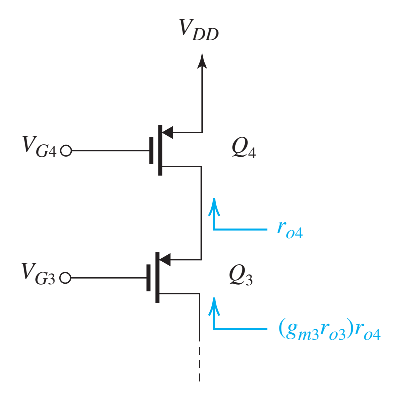


Figure 8.32 Employing a cascode transistor Q_3 to raise the output resistance of the current source Q_4.

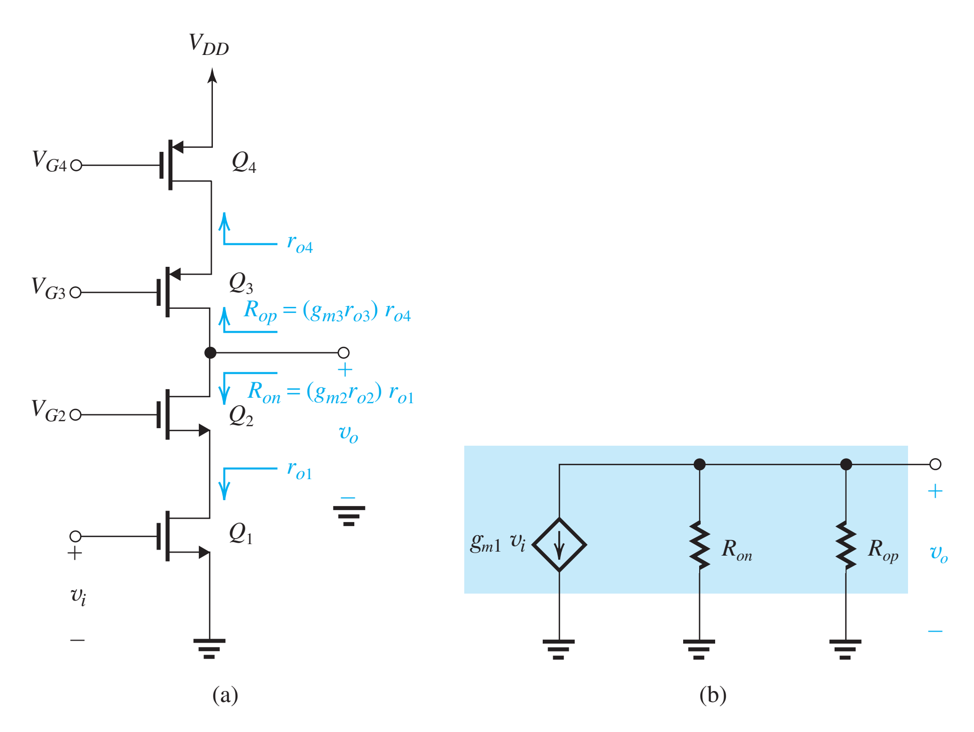


Figure 8.33 A cascode amplifier with a cascode current-source load.

output resistance for the cascode current source of

$$R_o = (g_{m3}r_{o3})r_{o4} \tag{8.77}$$

Combining a cascode amplifier with a cascode current source results in the circuit of Fig. 8.33(a). The equivalent circuit at the output side is shown in Fig. 8.33(b), from which the voltage gain can be easily found as

$$A_v = \frac{v_o}{v_i} = -g_{m1}\left[R_{on} \| R_{op}\right]$$

Thus,

$$A_v = -g_{m1}\{[(g_{m2}r_{o2})r_{o1}] \| [(g_{m3}r_{o3})r_{o4}]\} \tag{8.78}$$

For the case in which all transistors are identical,

$$A_v = -\frac{1}{2}(g_m r_o)^2 = -\frac{1}{2}A_0^2 \tag{8.79}$$

By comparison to the gain expression in Eq. (8.50), we see that using the cascode configuration for both the amplifying transistor and the current-source load transistor results in an increase in the magnitude of gain by a factor equal to A_0.

Example 8.5

It is required to design the cascode current source of Fig. 8.32 to provide a current of 100 μA and an output resistance of 500 kΩ. Assume the availability of a 0.18-μm CMOS technology for which $V_{DD} = 1.8$ V, $V_{tp} = -0.5$ V, $\mu_p C_{ox} = 90\ \mu\text{A/V}^2$, and $V_A' = -5$ V/μm. Use $|V_{OV}| = 0.3$ V and determine L and W/L for each transistor, and the values of the bias voltages V_{G3} and V_{G4}.

Solution

The output resistance R_o is given by

$$R_o = (g_{m3} r_{o3}) r_{o4}$$

Assuming Q_3 and Q_4 are identical,

$$R_o = (g_m r_o) r_o$$
$$= \frac{|V_A|}{|V_{OV}|/2} \times \frac{|V_A|}{I_D}$$

Using $|V_{OV}| = 0.3$ V, we write

$$500\ \text{k}\Omega = \frac{|V_A|}{0.15} \times \frac{|V_A|}{0.1\ \text{mA}}$$

Thus we require

$$|V_A| = 2.74\ \text{V}$$

Now, since $|V_A| = |V_A'|\, L$ we need to use a channel length of

$$L = \frac{2.74}{5} = 0.55\ \mu\text{m}$$

which is about three times the minimum channel length. With $|V_t| = 0.5$ V and $|V_{OV}| = 0.3$ V,

$$V_{SG4} = 0.5 + 0.3 = 0.8\ \text{V}$$

and thus,

$$V_{G4} = 1.8 - 0.8 = 1.0\ \text{V}$$

To allow for the largest possible signal swing at the output terminal, we shall use the minimum required voltage across Q_4, namely, $|V_{OV}|$ or 0.3 V. Thus,

$$V_{D4} = 1.8 - 0.3 = 1.5\ \text{V}$$

Since the two transistors are identical and are carrying equal currents,

$$V_{SG3} = V_{SG4} = 0.8\ \text{V}$$

Thus,

$$V_{G3} = 1.5 - 0.8 = +0.7\ \text{V}$$

Example 8.5 *continued*

We note that the maximum voltage allowed at the output terminal of the current source will be constrained by the need to allow a minimum voltage of $|V_{OV}|$ across Q_3; thus;

$$v_{D3\max} = 1.5 - 0.3 = +1.2 \text{ V}$$

To determine the required W/L ratios of Q_3 and Q_4, we use

$$I_D = \frac{1}{2}(\mu_p C_{ox})\left(\frac{W}{L}\right)|V_{OV}|^2\left(1 + \frac{V_{SD}}{|V_A|}\right)$$

$$100 = \frac{1}{2} \times 90 \times \left(\frac{W}{L}\right) \times 0.3^2\left(1 + \frac{0.3}{2.74}\right)$$

which yields

$$\frac{W}{L} = 22.3$$

EXERCISES

D8.18 If in Example 8.5, L of each of Q_3 and Q_4 is halved while W/L is changed to allow I_D and V_{OV} to remain unchanged, find the new values of R_o and W/L. [*Hint*: In computing the required (W/L), note that $|V_A|$ has changed.]
Ans. 125 kΩ; 20.3

8.19 Consider the cascode amplifier of Fig. 8.33 with the dc component at the input, $V_I = 0.7$ V, $V_{G2} = 1.0$ V, $V_{G3} = 0.8$ V, $V_{G4} = 1.1$ V, and $V_{DD} = 1.8$ V. If all devices are matched (i.e., if $k_{n1} = k_{n2} = k_{p3} = k_{p4}$), and have equal $|V_t|$ of 0.5 V, what is the overdrive voltage at which the four transistors are operating? What is the allowable voltage range at the output?
Ans. 0.2 V; 0.5 V to 1.3 V

8.20 The cascode amplifier in Fig. 8.33 is operated at a current of 0.2 mA with all devices operating at $|V_{OV}| = 0.2$ V. All devices have $|V_A| = 2$ V. Find g_{m1}, the output resistance of the amplifier, R_{on}, and the output resistance of the current source, R_{op}. Also find the overall output resistance and the voltage gain realized.
Ans. 2 mA/V; 200 kΩ, 200 kΩ; 100 kΩ; −200 V/V

8.5.3 Distribution of Voltage Gain in a Cascode Amplifier

It is often useful to know how much of the overall voltage gain of a cascode amplifier is realized in each of its two stages: the CS stage Q_1, and the CG stage Q_2. For this purpose, consider the cascode amplifier shown in Fig. 8.34(a). Here, for generality we have included

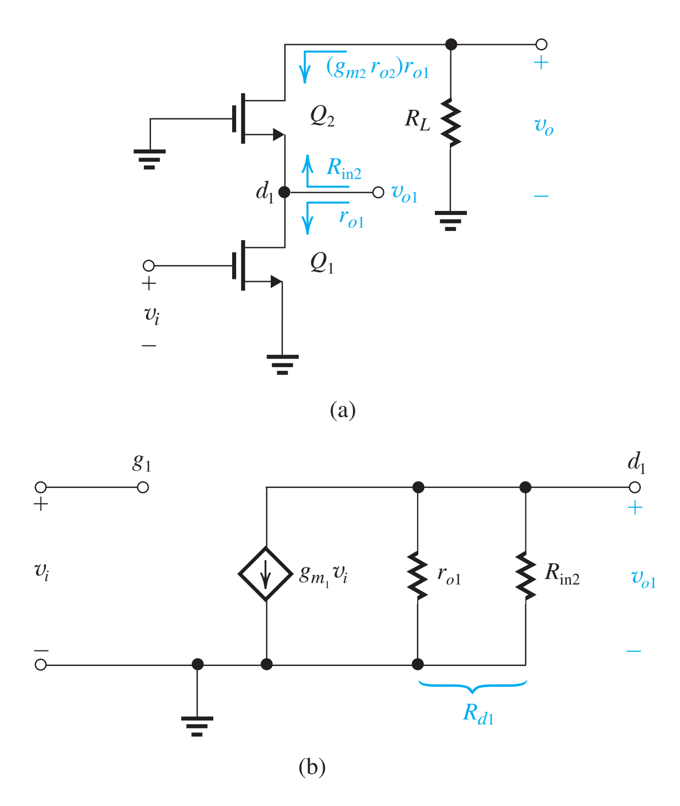


Figure 8.34 (**a**) The cascode amplifier with a load resistance R_L. Only signal quantities are shown. (**b**) Determining v_{o1}.

a load resistance R_L, which represents the output resistance of the current-source load plus any additional resistance that may be connected to the output node. The voltage gain A_v of the amplifier can be found as

$$A_v = -g_{m1}(R_o \| R_L)$$

Thus,

$$A_v = -g_{m1}(g_{m2}r_{o2}r_{o1} \| R_L) \quad (8.80)$$

The overall gain A_v can be expressed as the product of the voltage gains of Q_1 and Q_2 as

$$A_v = A_{v1}A_{v2} = \left(\frac{v_{o1}}{v_i}\right)\left(\frac{v_o}{v_{o1}}\right) \quad (8.81)$$

To obtain $A_{v1} \equiv v_{o1}/v_i$ we need to find the total resistance between the drain of Q_1 and ground. Referring to Fig. 8.34(b) and denoting this resistance R_{d1}, we can express A_{v1} as

$$A_{v1} = \frac{v_{o1}}{v_i} = -g_{m1}R_{d1} \quad (8.82)$$

Observe that R_{d1} is the parallel equivalent of r_{o1} and R_{in2}, where R_{in2} is the input resistance of the CG transistor Q_2. From Eq. (8.54), we can write

$$R_{in2} \simeq \frac{R_L}{g_{m2}r_{o2}} + \frac{1}{g_{m2}} \quad (8.83)$$

Table 8.1 Gain Distribution in the MOS Cascode Amplifier for Various Values of R_L

Case	R_L	R_{in2}	R_{d1}	A_{v1}	A_{v2}	A_v
1	∞	∞	r_o	$-g_m r_o$	$g_m r_o$	$-(g_m r_o)^2$
2	$(g_m r_o) r_o$	r_o	$r_o/2$	$-\frac{1}{2}(g_m r_o)$	$g_m r_o$	$-\frac{1}{2}(g_m r_o)^2$
3	r_o	$\frac{2}{g_m}$	$\frac{2}{g_m}$	-2	$\frac{1}{2}(g_m r_o)$	$-(g_m r_o)$
4	0	$\frac{1}{g_m}$	$\frac{1}{g_m}$	-1	0	0

Now we can obtain R_{d1} as

$$R_{d1} = r_{o1} \,\|\, R_{in2} \tag{8.84}$$

and A_{v1} as

$$A_{v1} = -g_{m1} R_{d1} = -g_{m1}(r_{o1} \,\|\, R_{in2}) \tag{8.85}$$

Finally, we can obtain A_{v2} by dividing the total gain A_v given by Eq. (8.80) by A_{v1}. To provide insight into the effect of the value of R_L on the overall gain of the cascode as well as on how this gain is distributed among the two stages of the cascode amplifier, we provide in Table 8.1 approximate values for the case $r_{o1} = r_{o2} = r_o$ and for four different values of R_L: (1) $R_L = \infty$, obtained with an ideal current-source load; (2) $R_L = (g_m r_o) r_o$, obtained with a cascode current-source load; (3) $R_L = r_o$, obtained with a simple current-source load; and (4) for completeness, $R_L = 0$, that is, a signal short circuit at the output.

Observe that while case 1 represents an idealized situation, it is useful in that it provides the theoretical maximum voltage gain achievable in a MOS cascode amplifier. Case 2, which assumes a cascode current-source load with an output resistance equal to that of the cascode amplifier, provides a realistic estimate of the gain achieved if one aims to maximize the realized gain. In certain situations, however, that is not our objective. This point is important, for as we shall see in Chapter 10, there is an entirely different application of the cascode amplifier: namely, to obtain wideband amplification by extending the upper-3-dB frequency f_H. As will be seen, for such an application one opts for the situation represented by case 3, where the gain achieved in the CS amplifier is only −2 V/V, and of course the overall gain is now only $-(g_m r_o)$. However, as will be seen in Chapter 10, this trade-off of the overall gain to obtain extended bandwidth is in some cases a good bargain!

EXERCISE

8.21 Consider a cascode amplifier for which the CS and CG transistors are identical and are biased to operate at $I_D = 0.1$ mA with $V_{OV} = 0.2$ V. Also let $V_A = 2$ V. Find A_{v1}, A_{v2}, and A_v for two cases: (a) $R_L = 20$ kΩ and (b) $R_L = 400$ kΩ.
Ans. (a) –1.82 V/V, 10.5 V/V, –19.0 V/V; (b) –10.2 V/V, 19.6 V/V, –200 V/V

8.5.4 Double Cascoding

If a still higher output resistance and correspondingly higher gain are required, it is possible to add another level of cascoding, as illustrated in Fig. 8.35. Observe that Q_3 is the second cascode transistor, and it raises the output resistance by $(g_{m3}r_{o3})$. For the case of identical transistors, the output resistance will be $(g_m r_o)^2 r_o$ and the voltage gain, assuming an ideal current-source load, will be $(g_m r_o)^3$ or A_0^3. Of course, we have to generate another dc bias voltage for the second cascode transistor, Q_3.

A drawback of double cascoding is that an additional transistor is now stacked between the power-supply rails. Furthermore, to realize the advantage of double cascoding, the current-source load will also need to use double cascoding with an additional transistor. Since for proper operation each transistor needs a certain minimum v_{DS} (at least equal to V_{OV}), and recalling that modern MOS technology utilizes power supplies in the range of 1 V to 2 V, we see that there is a limit on the number of transistors in a cascode stack.

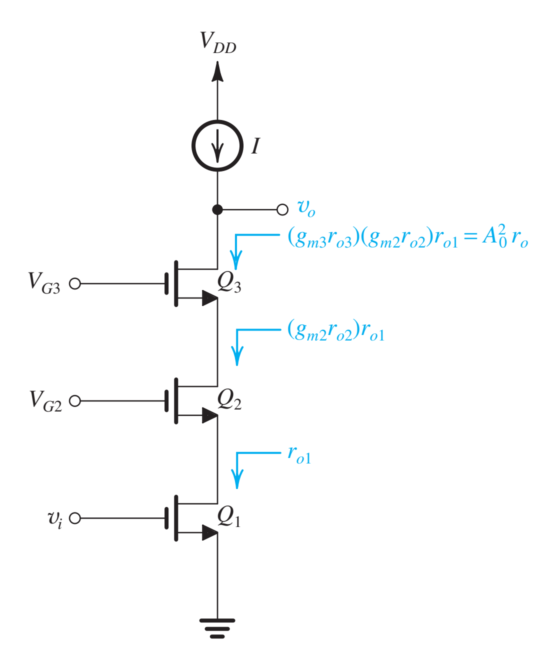


Figure 8.35 Double cascoding.

8.5.5 The Folded Cascode

To avoid the problem of stacking a large number of transistors across a low-voltage power supply, one can use a PMOS transistor for the cascode device, as shown in Fig. 8.36. Here, as before, the NMOS transistor Q_1 is operating in the CS configuration, but the CG stage is implemented using the PMOS transistor Q_2. An additional current source I_2 is needed to bias Q_2 and provide it with its active load. Note that Q_1 is now operating at a bias current of $(I_1 - I_2)$. Finally, a dc voltage V_{G2} is needed to provide an appropriate dc level for the gate of the cascode transistor Q_2. Its value has to be selected so that Q_2 and Q_1 operate in the saturation region.

The small-signal operation of the circuit in Fig. 8.36 is similar to that of the NMOS cascode. The difference here is that the signal current $g_m v_i$ is *folded down* and made to flow

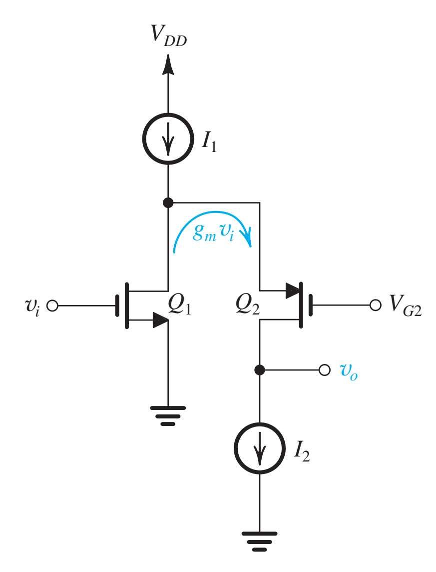


Figure 8.36 The folded cascode.

into the source terminal of Q_2, which gives the circuit the name **folded cascode**.[3] The folded cascode is a very popular building block in CMOS amplifiers.

EXERCISE

D8.22 Consider the folded-cascode amplifier of Fig. 8.36 for the following case: $V_{DD} = 1.8$ V, $k_n' = 4k_p'$, and $V_{tn} = -V_{tp} = 0.5$ V. To operate Q_1 and Q_2 at equal bias currents I, $I_1 = 2I$ and $I_2 = I$. While current source I_1 is implemented using the simple circuit studied in Section 8.2, current source I_2 is realized using a cascoded circuit (i.e., the NMOS version of the circuit in Fig. 8.32). The transistor W/L ratios are selected so that each operates at an overdrive voltage of 0.2 V.

(a) What must the relationship of $(W/L)_2$ to $(W/L)_1$ be?

(b) What is the minimum dc voltage required across current source I_1 for proper operation? Now, if a 0.1-V peak-to-peak signal swing is to be allowed at the drain of Q_1, what is the highest dc bias voltage that can be used at that node?

(c) What is the value of V_{SG} of Q_2, and hence what is the largest value to which V_{G2} can be set?

(d) What is the minimum dc voltage required across current-source I_2 for proper operation?

(e) Given the results of (c) and (d), what is the allowable range of signal swing at the output?

Ans. (a) $(W/L)_2 = 4\ (W/L)_1$; (b) 0.2 V, 1.55 V; (c) 0.7 V, 0.85 V; (d) 0.4 V; (e) 0.4 V to 1.35 V

[3]The circuit itself can be thought of as having been folded. In this same vein, the regular cascode is sometimes referred to as a **telescopic cascode** because the stacking of transistors resembles the extension of a telescope.

8.5.6 The BJT Cascode

Figure 8.37(a) shows the BJT cascode amplifier with an ideal current-source load. Voltage V_{B2} is a dc bias voltage for the CB cascode transistor Q_2. The circuit is very similar to the MOS cascode, and the small-signal analysis will follow in a parallel fashion. First, note that the input resistance of the bipolar cascode amplifier is finite,

$$R_{\text{in}} = r_{\pi 1} \tag{8.86}$$

Second, recall that the current signal in the collector of Q_2 will be approximately equal to $g_{m1} v_i$. Thus, the equivalent circuit of the output of the cascode amplifier will be that shown in Fig. 8.37(b). To obtain R_o we use the formula in Eq. (8.68) and note that the resistance R_e in the emitter of Q_2 is r_{o1}, thus

$$R_o \simeq r_{o2} + (g_{m2} r_{o2})(r_{o1} \parallel r_{\pi 2}) \tag{8.87}$$

Since $g_{m2}(r_{o1} \parallel r_{\pi 2}) \gg 1$, we can neglect the first term on the right-hand side of Eq. (8.87),

$$R_o \simeq (g_{m2} r_{o2})(r_{o1} \parallel r_{\pi 2}) \tag{8.88}$$

This result is similar but certainly *not identical* to that for the MOS cascode. Here, because of the finite β of the BJT, we have $r_{\pi 2}$ appearing in parallel with r_{o1}. This poses a very significant constraint on R_o of the BJT cascode. Specifically, because $(r_{o1} \parallel r_{\pi 2})$ will always be lower than $r_{\pi 2}$, it follows that the maximum possible value of R_o is

$$\begin{aligned} R_o\big|_{\text{max}} &= g_{m2} r_{o2} r_{\pi 2} \\ &= (g_{m2} r_{\pi 2}) r_{o2} = \beta_2 r_{o2} \end{aligned} \tag{8.89}$$

Thus the maximum output resistance realizable by cascoding is $\beta_2 r_{o2}$. This means that unlike the MOS case, double cascoding with a BJT would not be useful.

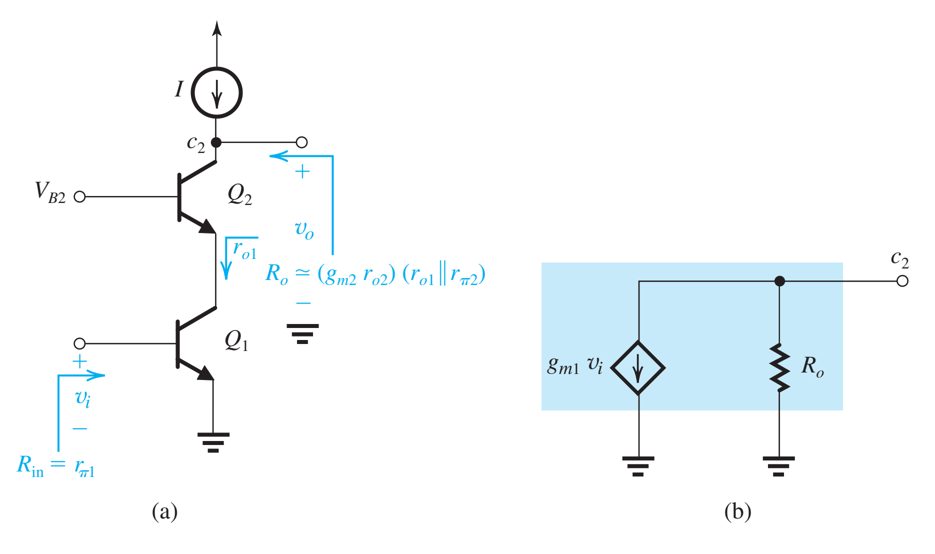


Figure 8.37 (**a**) A BJT cascode amplifier with an ideal current-source load; (**b**) small-signal, equivalent-circuit representation of the output of the cascode amplifier.

The open-circuit voltage gain of the bipolar cascode can be found using the equivalent circuit of Fig. 8.37(b) as

$$A_{vo} = \frac{v_o}{v_i} = -g_{m1}R_o$$

Thus,

$$A_{vo} = -g_{m1}(g_{m2}r_{o2})(r_{o1} \| r_{\pi 2}) \tag{8.90}$$

For the case $g_{m1} = g_{m2}$, $r_{o1} = r_{o2}$,

$$A_{vo} = -(g_m r_o)[g_m(r_o \| r_\pi)] \tag{8.91}$$

which will be less than $(g_m r_o)^2$ in magnitude. In fact, the maximum possible gain magnitude is obtained when $r_o \gg r_\pi$ and is given by

$$|A_{vo}|_{\max} = \beta g_m r_o = \beta A_0 \tag{8.92}$$

Finally, we note that to be able to realize gains approaching this level, the current-source load must also be cascoded. Figure 8.38 shows a cascode BJT amplifier with a cascode current-source load.

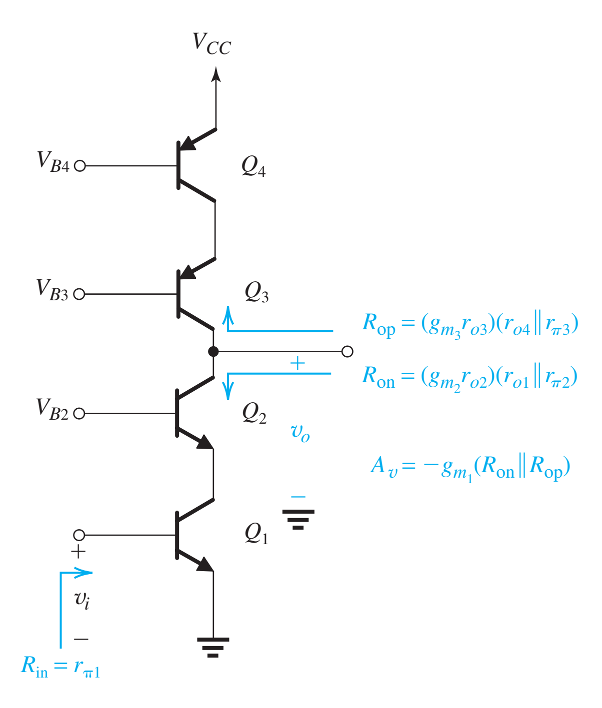


Figure 8.38 A BJT cascode amplifier with a cascode current source.

EXERCISES

8.23 Find an expression for the maximum voltage gain achieved in the amplifier of Fig. 8.38.
Ans. $|A_{vmax}| = g_{m1}(\beta_2 r_{o2} \| \beta_3 r_{o3})$

8.24 Consider the BJT cascode amplifier of Fig. 8.38 when biased at a current of 0.2 mA. Assuming that *npn* transistors have $\beta = 100$ and $V_A = 5$ V and that *pnp* transistors have $\beta = 50$ and $|V_A| = 4$ V, find R_{on}, R_{op}, and A_v. Also use the result of Exercise 8.23 to determine the maximum achievable gain.
Ans. 1.67 MΩ; 0.762 MΩ; −4186 V/V; −5714 V/V

8.6 Current-Mirror Circuits with Improved Performance



As we have seen throughout this chapter, current sources play a major role in the design of IC amplifiers: The constant-current source is used both in biasing and as active load. Simple forms of both MOS and bipolar current sources and, more generally, current mirrors were studied in Section 8.2. The need to improve the characteristics of the simple sources and mirrors has already been demonstrated.

Specifically, three performance parameters need to be addressed:

1. The accuracy of the current transfer ratio of the mirror. For bipolar mirrors, this parameter is primarily affected by the transistor β. For both bipolar and MOS mirrors, the Early effect affects the current transfer ratio.
2. The output resistance, R_o. The need to increase the output resistance of current sources is motivated by the need to increase the voltage gain achievable in an amplifier stage. While simple bipolar and MOS mirrors have output resistances equal to r_o, cascoding can be used to increase the output resistance.
3. The minimum dc voltage required across the current source. The need to keep this voltage as small as possible stems from the low dc voltage supplies employed in modern IC technologies. Simple BJT and MOS sources can operate with dc voltages in the range of 0.2 V to 0.3 V. More elaborate mirror circuits usually require higher voltages.

In this section we study MOS and bipolar current mirrors that feature improvements in one or more of these characteristics.

8.6.1 Cascode MOS Mirrors

The use of cascoding in the design of current sources was presented in Section 8.5. Figure 8.39 shows the basic cascode current mirror. Observe that in addition to the diode-connected transistor Q_1, which forms the basic mirror Q_1–Q_2, another diode-connected transistor, Q_4, is used to provide a suitable bias voltage for the gate of the cascode transistor Q_3. To determine the output resistance of the cascode mirror at the drain of Q_3, we assume that the voltages

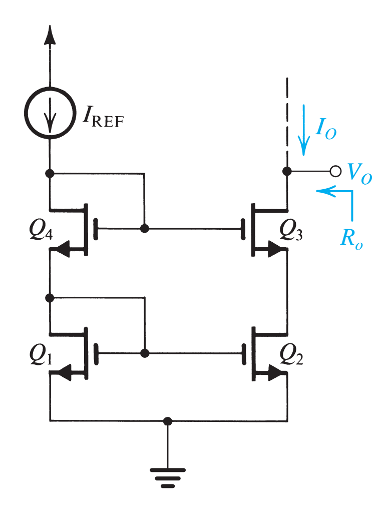


Figure 8.39 A cascode MOS current mirror.

across Q_1 and Q_4 are constant, and thus the signal voltages at the gates of Q_2 and Q_3 will be zero. Thus R_o will be that of the cascode current source formed by Q_2 and Q_3,

$$R_o \simeq g_{m3} r_{o3} r_{o2} \tag{8.93}$$

Thus, as expected, cascoding raises the output resistance of the current source by the factor $(g_{m3} r_{o3})$, which is the intrinsic gain of the cascode transistor.

A drawback of the cascode current mirror is that it consumes a relatively large portion of the steadily shrinking supply voltage V_{DD}. While the simple MOS mirror operates properly with a voltage as low as V_{OV} across its output transistor, the cascode circuit of Fig. 8.39 requires a minimum voltage of $V_t + 2V_{OV}$. This is because the gate of Q_3 is at $2V_{GS} = 2V_t + 2V_{OV}$. Thus the minimum voltage required across the output of the cascode mirror is 1 V or so. This obviously limits the signal swing at the output of the mirror (i.e., at the output of the amplifier that utilizes this current source as a load). In Chapter 13 we shall study a wide-swing cascode mirror.

EXERCISE

8.25 For a cascode MOS mirror utilizing devices with $V_t = 0.5$ V, $\mu_n C_{ox} = 387$ μA/V^2, $V_A' = 5$ V/μm, $W/L = 3.6$ μm/0.36 μm, and $I_{REF} = 100$ μA, find the minimum voltage required at the output and the output resistance.
Ans. 0.95 V; 285 kΩ

8.6.2 The Wilson Current Mirror

A simple but ingenious modification of the basic bipolar mirror results in both reducing the β dependence and increasing the output resistance. The resulting circuit, known as the **Wilson mirror** after its inventor George Wilson, an IC design engineer working for Tektronix, is shown in Fig. 8.40(a). The analysis to determine the effect of finite β on the current transfer

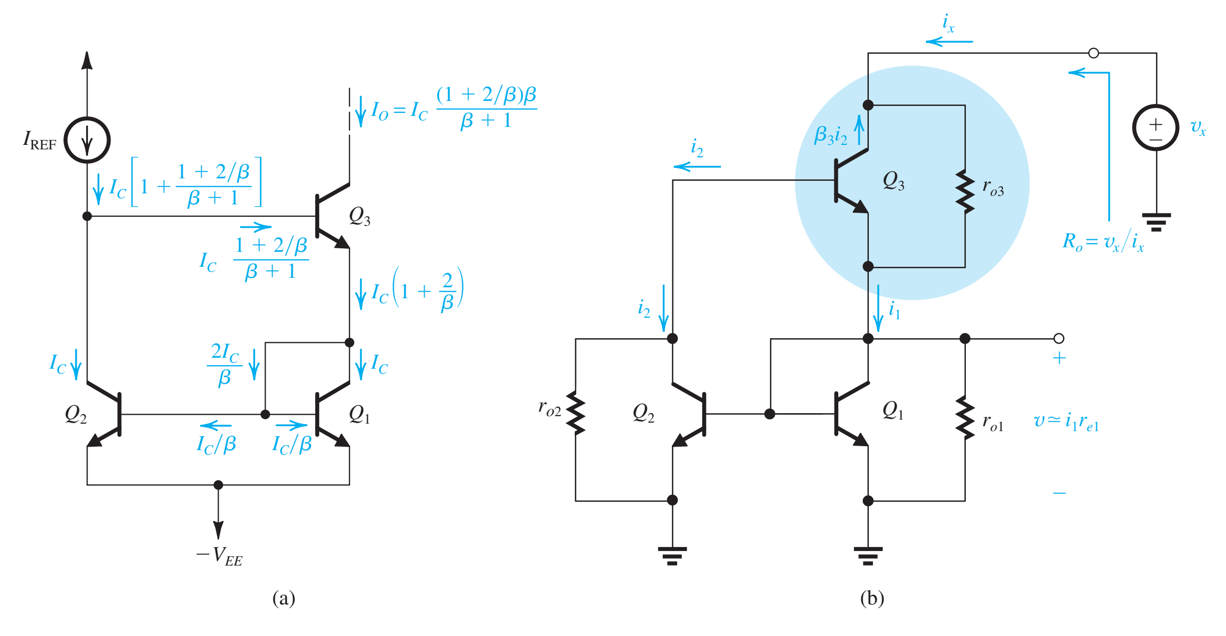


Figure 8.40 The Wilson bipolar current mirror: (**a**) circuit showing analysis to determine the current transfer ratio; (**b**) determining the output resistance.

ratio is shown in Fig. 8.40(a), from which we can write

$$\frac{I_O}{I_{REF}} = \frac{I_C\left(1+\frac{2}{\beta}\right)\beta\Big/(\beta+1)}{I_C\left[1+\left(1+\frac{2}{\beta}\right)\Big/(\beta+1)\right]}$$

$$= \frac{\beta+2}{\beta+1+\frac{\beta+2}{\beta}} = \frac{\beta+2}{\beta+2+\frac{2}{\beta}}$$

$$= \frac{1}{1+\frac{2}{\beta(\beta+2)}}$$

$$\simeq \frac{1}{1+2/\beta^2} \qquad (8.94)$$

which is much less dependent on β than in the case of the simple current mirror.

This analysis assumes that Q_1 and Q_2 conduct equal collector currents. There is, however, a slight problem with this assumption: The collector-to-emitter voltages of Q_1 and Q_2 are not equal, which introduces a current offset or a systematic error. The problem can be solved by adding a diode-connected transistor in series with the collector of Q_2, as we shall shortly show for the MOS version.

To determine the output resistance of the Wilson mirror, we set $I_{REF} = 0$ and apply a test voltage v_x to the output node, as shown in Fig. 8.40(b). Our purpose is to determine the current i_x and hence R_o as

$$R_o = v_x/i_x$$

Rather than replacing each transistor with its hybrid-π model, we shall do the analysis directly on the circuit diagram. For this purpose, we have "pulled r_o out" of each transistor and shown it separately.

Observe that transistor Q_3, viewed as a supernode (highlighted in color), has a current i_x entering it and two currents i_1 and i_2 exiting it; thus,

$$i_1 + i_2 = i_x$$

Next note that the action of current mirror Q_1–Q_2 forces i_2 to be approximately equal to i_1; thus,

$$i_2 \simeq i_1 = i_x/2$$

Current i_2 flows into the base of Q_3 and thus gives rise to a collector current $\beta_3 i_2$ in the direction indicated. We are now in a position to write a node equation at the collector of Q_3 and thus determine the current through r_{o3} as $i_x + \beta_3 i_2 = i_x + \beta_3(i_x/2) = i_x(\beta_3/2 + 1)$. Finally, we can express the voltage between the collector of Q_3 and ground as the sum of the voltage drop across r_{o3} and the voltage v across Q_1,

$$\begin{aligned} v_x &= i_x\left(\frac{\beta_3}{2} + 1\right)r_{o3} + i_1 r_{e1} \\ &= i_x\left(\frac{\beta_3}{2} + 1\right)r_{o3} + \left(\frac{i_x}{2}\right)r_{e1} \end{aligned}$$

Since $r_o \gg r_e$ and $\beta_3 \gg 2$

$$v_x \simeq i_x\left(\frac{\beta_3}{2}\right)r_{o3}$$

and

$$R_o = \beta_3 r_{o3}/2 \tag{8.95}$$

Thus the Wilson current mirror has an output resistance $(\frac{1}{2}\beta_3)$ times higher than that of Q_3 alone. This is a result of the negative feedback obtained by feeding the collector current of Q_2 (i_2) back to the base of Q_3. As can be seen from the above analysis, this feedback results in increasing the current through r_{o3} to approximately $\frac{1}{2}\beta_3 i_x$, and thus the voltage across r_{o3} and the output resistance increase by the same factor, $\frac{1}{2}\beta_3$. Finally, note that the factor $\frac{1}{2}$ is because only half of i_x is mirrored back to the base of Q_3.

The Wilson mirror is preferred over the cascode circuit because the latter has the same dependence on β as the simple mirror. However, like the cascode mirror, the Wilson mirror requires an additional V_{BE} drop for its operation; that is, for proper operation we must allow for 1 V or so across the Wilson mirror output.

EXERCISE

8.26 For $\beta = 100$ and $r_o = 100\,\text{k}\Omega$, contrast the Wilson mirror and the simple mirror by evaluating the transfer-ratio error due to finite β, and the output resistance.
Ans. Transfer-ratio error: 0.02% for Wilson as opposed to 2% for the simple circuit; $R_o = 5\,\text{M}\Omega$ for Wilson compared to 100 kΩ for the simple circuit

8.6.3 The Wilson MOS Mirror

Figure 8.41(a) shows the MOS version of the Wilson mirror. Obviously there is no β error to reduce here, and the advantage of the MOS Wilson lies in its enhanced output resistance.

To determine the output resistance of the Wilson MOS mirror, we set $I_{REF} = 0$, and apply a test voltage v_x to the output node, as shown in Fig. 8.41(b). Our purpose is to determine the current i_x and hence R_o as

$$R_o = v_x/i_x$$

Rather than replacing each transistor with its hybrid-π equivalent-circuit model, we shall perform the analysis directly on the circuit. For this purpose, we have "pulled r_o out" of each transistor and shown it separately.

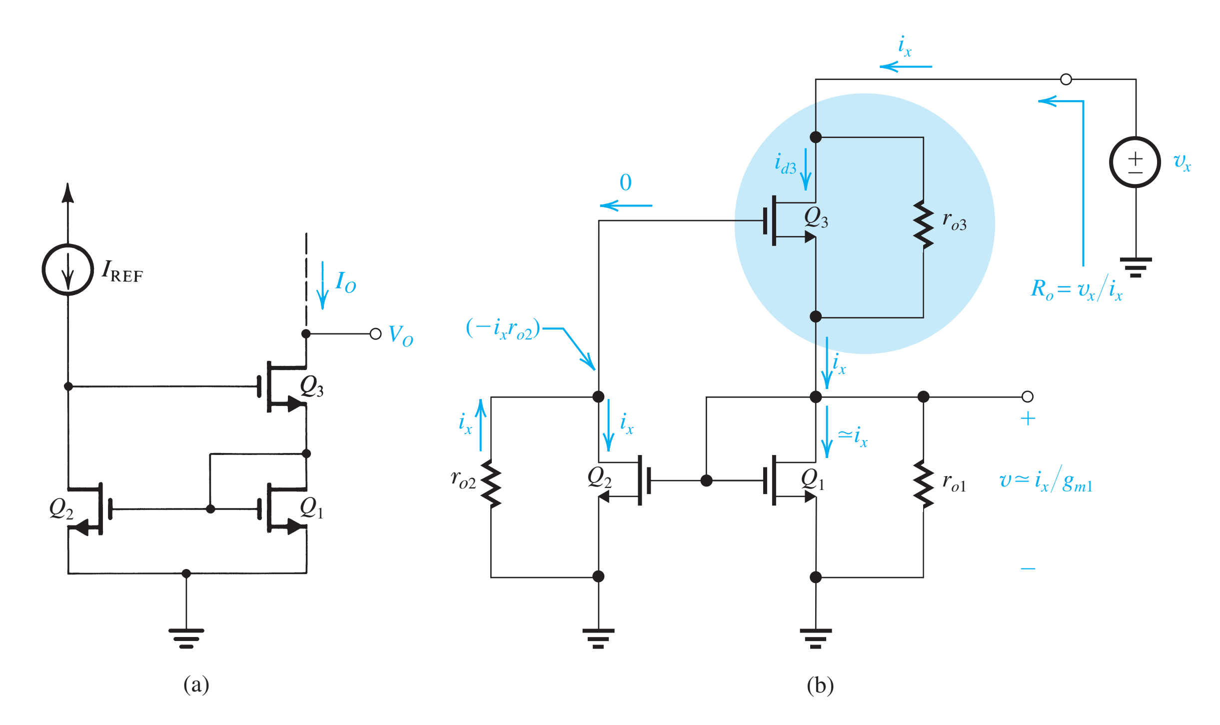
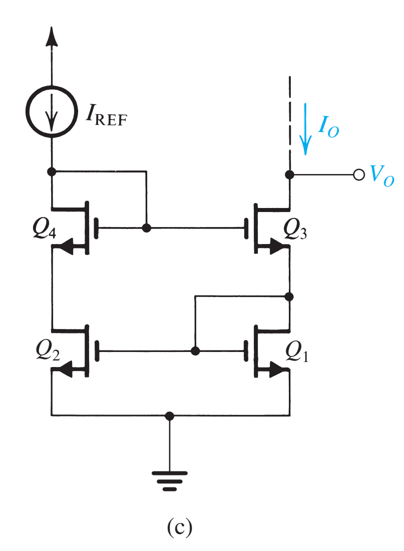


Figure 8.41 The Wilson MOS mirror: (**a**) circuit; (**b**) analysis to determine output resistance; (**c**) modified circuit.



(c)

Figure 8.41 *continued*

Observe that the current i_x that enters the drain of Q_3 must exit at its source. Thus the current that feeds the input side of the Q_1–Q_2 mirror is equal to i_x. Most of this current will flow in the drain proper of Q_1 (i.e., only a very small fraction flows through r_{o1}) and will give rise to a voltage $v \simeq i_x/g_{m1}$, where $1/g_{m1}$ is the approximate resistance of the diode-connected transistor Q_1. The current-mirror action of (Q_1, Q_2) forces a current equal to i_x to flow through the drain proper of Q_2. Now, since the current in the drain of Q_2 is forced (by the connection to the gate of Q_3) to be zero, all of i_x must flow through r_{o2}, resulting in a voltage $-i_x r_{o2}$. This is the voltage fed back to the gate of Q_3. The drain current of Q_3 can now be found as

$$\begin{aligned} i_{d3} &= g_{m3} v_{gs3} \\ &= g_{m3}\left(v_{g3} - v_{s3}\right) \\ &= g_{m3}(-i_x r_{o2} - i_x/g_{m1}) \\ &\simeq -g_{m3} r_{o2} i_x \end{aligned}$$

A node equation at the drain of Q_3 gives the current through r_{o3} as $(i_x - i_{d3}) = i_x + g_{m3} r_{o2} i_x \simeq g_{m3} r_{o2} i_x$. Finally, we can express v_x as the sum of the voltage drop across r_{o3} and the voltage v across Q_1,

$$\begin{aligned} v_x &= g_{m3} r_{o2} i_x r_{o3} + v \\ &= (g_{m3} r_{o3} r_{o2}) i_x + (i_x/g_{m1}) \\ &\simeq g_{m3} r_{o3} r_{o2} i_x \end{aligned}$$

and obtain



$$R_o = \frac{v_x}{i_x} = (g_{m3} r_{o3}) r_{o2} \tag{8.96}$$

Thus, the Wilson MOS mirror exhibits an increase of output resistance by a factor $(g_{m3} r_{o3})$, an identical result to that achieved in the cascode mirror. Here the increase in R_o, as demonstrated in the analysis above, is a result of the negative feedback obtained by connecting the drain of Q_2 to the gate of Q_3. Finally, to balance the two branches of the mirror and thus avoid the systematic current error resulting from the difference in V_{DS} between Q_1 and Q_2, the circuit can be modified as shown in Fig. 8.41(c).

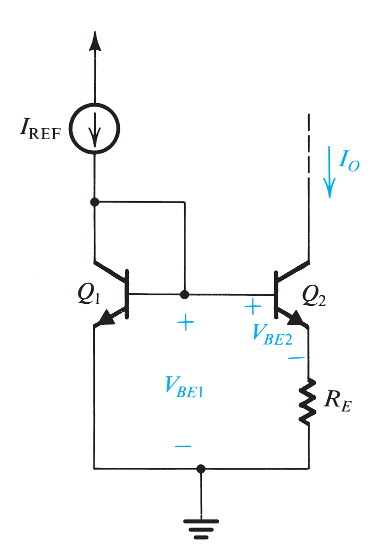


Figure 8.42 The Widlar current source.

8.6.4 The Widlar Current Source[4]

Our final current-source circuit, known as the **Widlar current source**, is shown in Fig. 8.42. It differs from the basic current-mirror circuit in an important way: A resistor R_E is included in the emitter lead of Q_2. Neglecting base currents we can write

$$V_{BE1} = V_T \ln\left(\frac{I_{REF}}{I_S}\right) \tag{8.97}$$

and

$$V_{BE2} = V_T \ln\left(\frac{I_O}{I_S}\right) \tag{8.98}$$

where we have assumed that Q_1 and Q_2 are matched devices. Combining Eqs. (8.97) and (8.98) gives

$$V_{BE1} - V_{BE2} = V_T \ln\left(\frac{I_{REF}}{I_O}\right) \tag{8.99}$$

But from the circuit we see that

$$V_{BE1} = V_{BE2} + I_O R_E \tag{8.100}$$

Thus,

$$I_O R_E = V_T \ln\left(\frac{I_{REF}}{I_O}\right) \tag{8.101}$$

The design and advantages of the Widlar current source are illustrated in the following example.

---

[4]Named after Robert Widlar, a pioneer in analog IC design.

Example 8.6

The two circuits for generating a constant current $I_O = 10\ \mu\text{A}$ shown in Fig. 8.43 operate from a 10-V supply. Determine the values of the required resistors, assuming that V_{BE} is 0.7 V at a current of 1 mA and neglecting the effect of finite β.

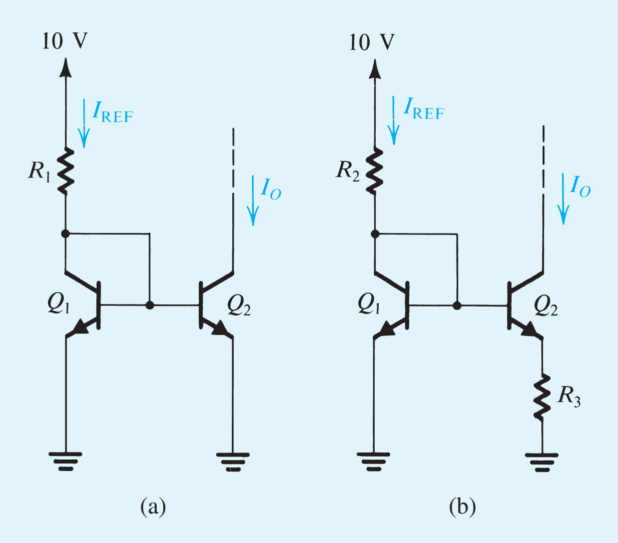


Figure 8.43 Circuits for Example 8.6.

Solution

For the basic current-source circuit in Fig. 8.43(a) we choose a value for R_1 to result in $I_{REF} = 10\ \mu\text{A}$. At this current, the voltage drop across Q_1 will be

$$V_{BE1} = 0.7 + V_T \ln\left(\frac{10\ \mu\text{A}}{1\text{ mA}}\right) = 0.58\text{ V}$$

Thus,

$$R_1 = \frac{10 - 0.58}{0.01} = 942\text{ k}\Omega$$

For the Widlar circuit in Fig. 8.43(b) we must first decide on a suitable value for I_{REF}. If we select $I_{REF} =$ 1 mA, then $V_{BE1} = 0.7$ V and R_2 is given by

$$R_2 = \frac{10 - 0.7}{1} = 9.3\text{ k}\Omega$$

The value of R_3 can be determined using Eq. (8.101) as follows:

$$10 \times 10^{-6} R_3 = 0.025 \ln\left(\frac{1\text{ mA}}{10\ \mu\text{A}}\right)$$

$$R_3 = 11.5\text{ k}\Omega$$

From the above example we observe that using the Widlar circuit allows the generation of a small constant current using relatively small resistors. This is an important advantage that results in considerable savings in chip area. In fact the circuit of Fig. 8.43(a), requiring a 942-kΩ resistance, is totally impractical for implementation in IC form because of the very high value of resistor R_1.

Another important characteristic of the Widlar current source is that its output resistance is high. The increase in the output resistance, above that achieved in the basic current source, is due to the emitter-degeneration resistance R_E. To determine the output resistance of Q_2, we assume that since the base of Q_2 is connected to ground via the small resistance r_e of Q_1, the incremental voltage at the base will be small. Thus we can use the formula in Eq. (8.70) and adapt it for our purposes here as follows:

$$R_{\text{out}} \simeq [1 + g_m(R_E \| r_\pi)]r_o \tag{8.102}$$

Thus the output resistance is increased above r_o by a factor that can be significant.

EXERCISE

8.27 Find the output resistance of each of the two current sources designed in Example 8.6. Let $V_A = 100$ V and $\beta = 100$.
Ans. 10 MΩ; 54 MΩ

8.7 Some Useful Transistor Pairings



The cascode configuration studied in Section 8.5 combines CS and CG MOS transistors (CE and CB bipolar transistors) to great advantage. The key to the superior performance of the resulting combination is that the transistor pairing is done in a way that maximizes the advantages and minimizes the shortcomings of each of the two individual configurations. In this section we present a number of other such transistor pairings. In each case the transistor pair can be thought of as a compound device; thus the resulting amplifier may be considered as a single stage.

8.7.1 The CC–CE, CD–CS, and CD–CE Configurations

Figure 8.44(a) shows an amplifier formed by cascading a common-collector (emitter-follower) transistor Q_1 with a common-emitter transistor Q_2. This circuit has two main advantages over the CE amplifier. First, the emitter follower increases the input resistance by a factor equal to $(\beta_1 + 1)$. As a result, the overall voltage gain is increased, especially if the resistance of the signal source is large. Second, it will be shown in Chapter 10 that the CC–CE amplifier can exhibit much wider bandwidth than that obtained with the CE amplifier.

The MOS counterpart of the CC–CE amplifier, namely, the CD–CS configuration, is shown in Fig. 8.44(b). Here, since the CS amplifier alone has an infinite input resistance, the

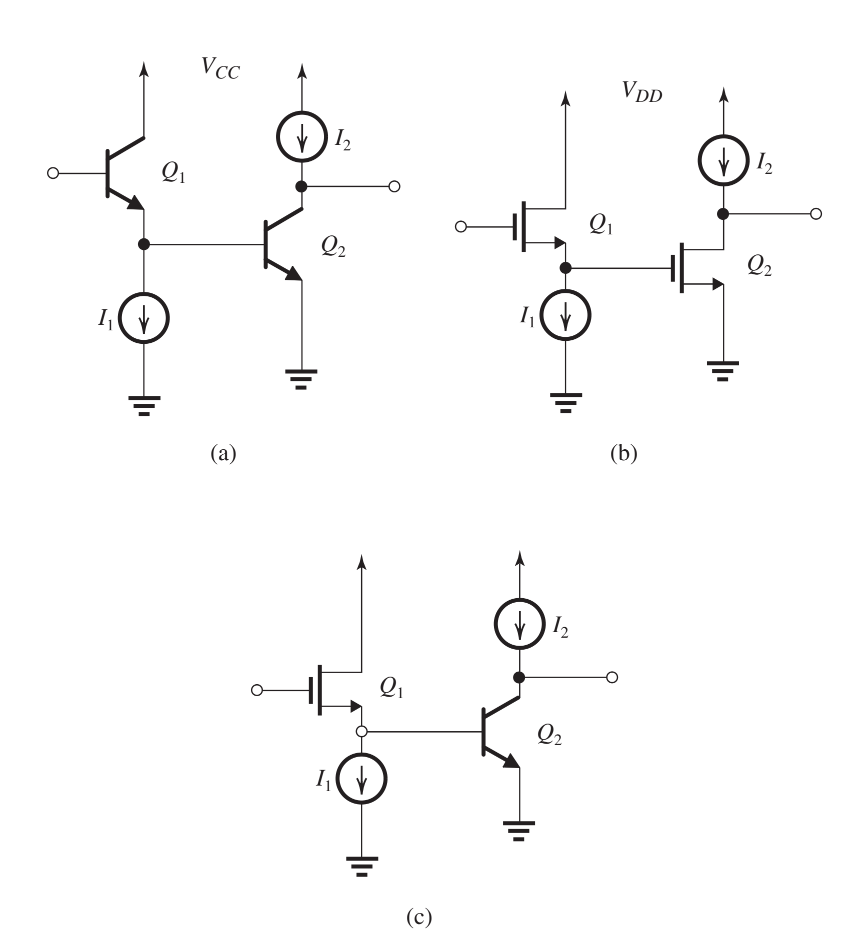


Figure 8.44 (**a**) CC–CE amplifier; (**b**) CD–CS amplifier; (**c**) CD–CE amplifier.

sole purpose for adding the source-follower stage is to increase the amplifier bandwidth, as will be seen in Chapter 10. Finally, Fig. 8.44(c) shows the BiCMOS version of this circuit type. Compared to the bipolar circuit in Fig. 8.44(a), the BiCMOS circuit has an infinite input resistance. Compared to the MOS circuit in Fig. 8.44(b), the BiCMOS circuit typically has a higher g_{m2}.

The IC Source Follower Since a number of the circuit configurations discussed in this section utilize an input source follower, we digress briefly to consider the IC source follower (the discrete-circuit source follower was studied in Section 7.3.6). Figure 8.45(a) shows a source follower formed by transistor Q_1 and biased by a constant-current supplied by the current mirror Q_2–Q_3. Observe that since the source of Q_1 cannot be connected to the body (which is at signal ground potential) a voltage signal v_{bs} develops between body and source and gives rise to a current signal $g_{mb}v_{bs}$, as indicated in the equivalent circuit in Fig. 8.45(b). The equivalent circuit shows also the output resistance r_{o3} of the bias current source Q_3, which acts as a load resistance for the follower Q_1.

An important observation to make from the equivalent circuit is that the controlled source ($g_{mb}v_{bs}$) appears across its control voltage v_{bs}. Thus we can use the source-absorption theorem (Appendix G) to replace the controlled source with a resistance $1/g_{mb}$. Next, note that the three

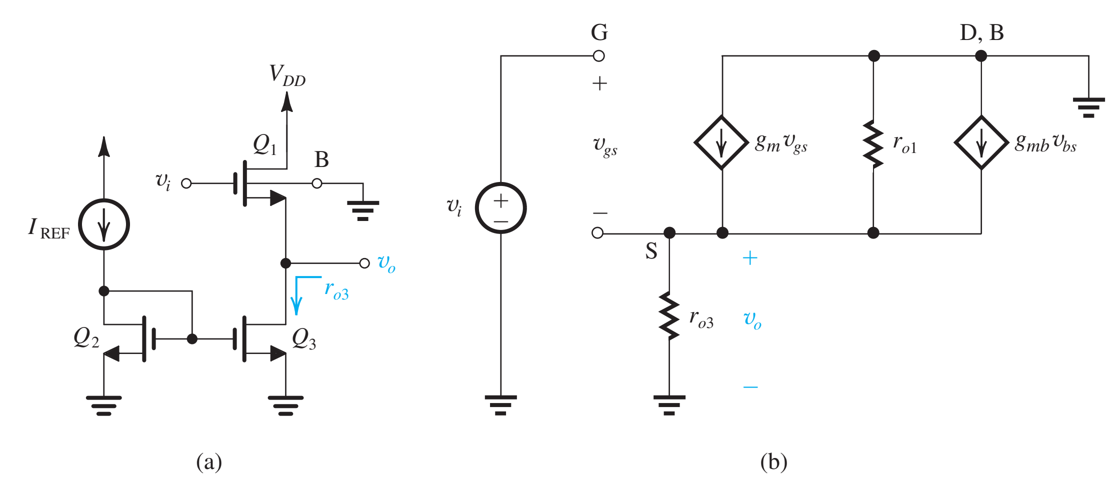


Figure 8.45 (**a**) A source follower biased with a current mirror Q_2–Q_3 and with the body terminal indicated. Note that the source cannot be connected to the body and thus the body effect should be taken into account. (**b**) Equivalent circuit.

resistances $1/g_{mb}$, r_{o1}, and r_{o3} appear in parallel between the source and ground. If we denote their parallel equivalent R_L, we can easily show that the voltage gain of the source follower is given by

$$\frac{v_o}{v_i} = \frac{R_L}{R_L + \dfrac{1}{g_m}} \tag{8.103}$$

where

$$R_L = r_{o1} \,\|\, r_{o3} \,\|\, \frac{1}{g_{mb}} \tag{8.104}$$

In cases where $\dfrac{1}{g_{mb}} \ll r_{o1}, r_{o3}$,

$$R_L \simeq \frac{1}{g_{mb}}$$

and

$$\frac{v_o}{v_i} \simeq \frac{g_{mb}}{g_m + g_{mb}} \tag{8.105}$$

Substituting for $g_{mb} = \chi g_m$ where $\chi = 0.1$ to 0.2,

$$\frac{v_o}{v_i} \simeq \frac{1}{1+\chi} \tag{8.106}$$

This is the maximum possible gain obtained from an IC source follower. The actual gain realized will usually be lower because of the effect of r_{o1} and r_{o3}.

EXERCISE

8.28 For the source follower in Fig. 8.45(a), let the bias current of Q_1 be 200 μA and assume Q_1 is operating at $V_{OV} = 0.2$ V. If $V_A = 5$ V and $\chi = 0.2$, find the voltage gain of the source follower.
Ans. 0.81 V/V

Example 8.7

For the CC–CE amplifier in Fig. 8.44(a) let $I_1 = I_2 = 1$ mA and assume identical transistors with $\beta = 100$. Find the input resistance R_{in} and the overall voltage gain obtained when the amplifier is fed with a signal source having $R_{sig} = 4\ \text{k}\Omega$ and loaded with a resistance $R_L = 4\ \text{k}\Omega$. Compare the results with those obtained with a common-emitter amplifier operating under the same conditions. Ignore r_o.

Solution

At an emitter current of 1 mA, Q_1 and Q_2 have

$$g_m = 40\ \text{mA/V}$$
$$r_e = 25\ \Omega$$
$$r_\pi = \frac{\beta}{g_m} = \frac{100}{40} = 2.5\ \text{k}\Omega$$

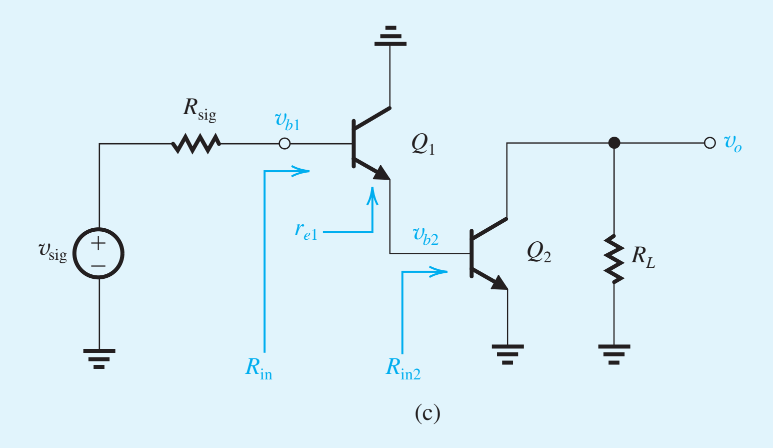


Figure 8.46 Circuit for Example 8.7.

Referring to Fig. 8.46 we can find

$$R_{in2} = r_{\pi 2} = 2.5\ \text{k}\Omega$$
$$R_{in} = (\beta_1 + 1)(r_{e1} + R_{in2})$$
$$= 101(0.025 + 2.5) = 255\ \text{k}\Omega$$
$$\frac{v_{b1}}{v_{sig}} = \frac{R_{in}}{R_{in} + R_{sig}} = \frac{255}{255 + 4} = 0.98\ \text{V/V}$$
$$\frac{v_{b2}}{v_{b1}} = \frac{R_{in2}}{R_{in2} + r_{e1}} = \frac{2.5}{2.5 + 0.025} = 0.99\ \text{V/V}$$
$$\frac{v_o}{v_{b2}} = -g_{m2}R_L = -40 \times 4 = -160\ \text{V/V}$$

Thus,

$$G_v = \frac{v_o}{v_{sig}} = -160 \times 0.99 \times 0.98 = -155 \text{ V/V}$$

For comparison, a CE amplifier operating under the same conditions will have

$$R_{in} = r_\pi = 2.5 \text{ k}\Omega$$

$$G_v = \frac{R_{in}}{R_{in} + R_{sig}}(-g_m R_L)$$

$$= \frac{2.5}{2.5 + 4}(-40 \times 4)$$

$$= -61.5 \text{ V/V}$$

EXERCISE

8.29 Repeat Example 8.7 for the CD–CE configuration of Fig. 8.44(c). Let $I_1 = I_2 = 1$ mA, $\beta_2 = 100$, $R_L = 4$ kΩ, and $k_{n1} = 8$ mA/V^2; neglect the body effect in Q_1 and r_o of both transistors. Find R_{in} and G_v when $R_{sig} = 4$ kΩ (as in Example 8.7) and $R_{sig} = 400$ kΩ. What would G_v of the CC–CE amplifier in Example 8.7 become for $R_{sig} = 400$ kΩ?
Ans. $R_{in} = \infty$; $G_v = -145.5$ V/V, independent of R_{sig}; -61.7 V/V

8.7.2 The Darlington Configuration[5]

Figure 8.47(a) shows a popular BJT circuit known as the **Darlington configuration**. It can be thought of as a variation of the CC–CE circuit with the collector of Q_1 connected to that of Q_2. Alternatively, the **Darlington pair** can be thought of as a composite transistor with $\beta = \beta_1 \beta_2$. It can therefore be used to implement a high-performance voltage follower, as illustrated in Fig. 8.47(b). Note that in this application the circuit can be considered as the cascade connection of two common-collector transistors (i.e., a CC–CC configuration).

Since the transistor β depends on the dc bias current, it is possible that Q_1 will be operating at a very low β, rendering the β-multiplication effect of the Darlington pair rather ineffective. A simple solution to this problem is to provide a bias current for Q_1, as shown in Fig. 8.47(c).

[5]Named after Sidney Darlington, a pioneer in filter design and transistor circuit design.

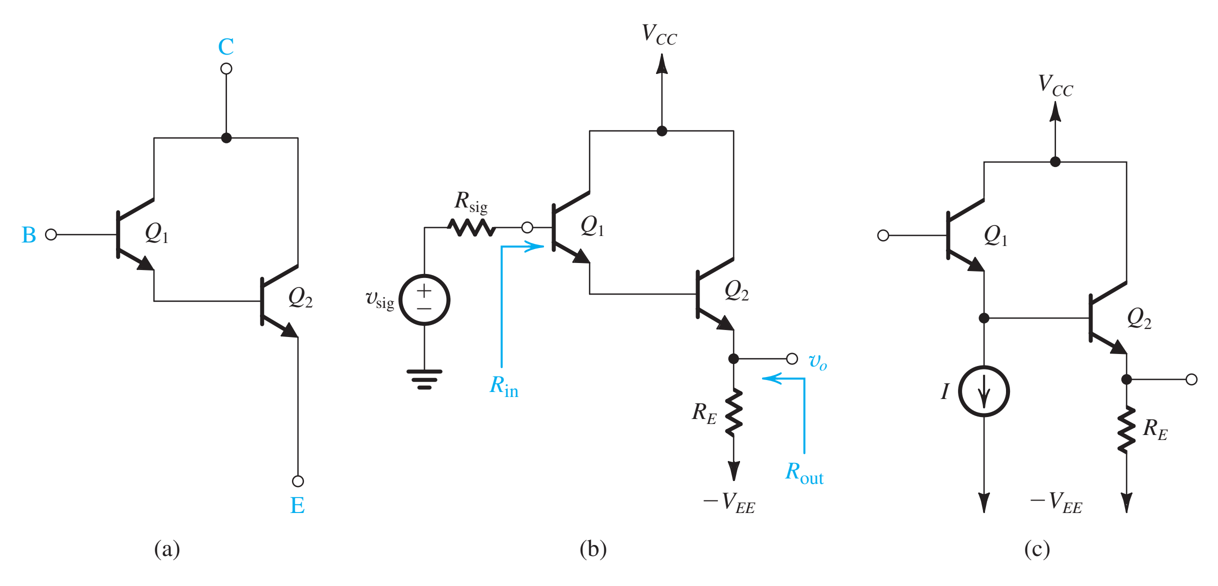


Figure 8.47 (**a**) The Darlington configuration; (**b**) voltage follower using the Darlington configuration; (**c**) the Darlington follower with a bias current I supplied to Q_1 to ensure that its β remains high.

EXERCISE

8.30 For the Darlington voltage follower in Fig. 8.47(b), show that:

$$R_{in} = (\beta_1 + 1)[r_{e1} + (\beta_2 + 1)(r_{e2} + R_E)]$$

$$R_{out} = R_E \parallel \left[r_{e2} + \frac{r_{e1} + [R_{sig}/(\beta_1 + 1)]}{\beta_2 + 1} \right]$$

$$\frac{v_o}{v_{sig}} = \frac{R_E}{R_E + r_{e2} + [r_{e1} + R_{sig}/(\beta_1 + 1)]/(\beta_2 + 1)}$$

Evaluate R_{in}, R_{out}, and v_o/v_{sig} for the case $I_{E2} = 5$ mA, $\beta_1 = \beta_2 = 100$, $R_E = 1$ kΩ, and $R_{sig} = 100$ kΩ.
Ans. 10.3 MΩ; 20 Ω; 0.98 V/V

8.7.3 The CC–CB and CD–CG Configurations

Cascading an emitter follower with a common-base amplifier, as shown in Fig. 8.48(a), results in a circuit with a low-frequency gain approximately equal to that of the CB but with the problem of the low input resistance of the CB solved by the buffering action of the CC stage. It will be shown in Chapter 10 that this circuit exhibits wider bandwidth than that obtained with a CE amplifier of the same gain. Note that the biasing current sources shown in Fig. 8.48(a) ensure that each of Q_1 and Q_2 is operating at a bias current I. We are not showing, however, how the dc voltage at the base of Q_1 is set, nor do we show the circuit that determines the

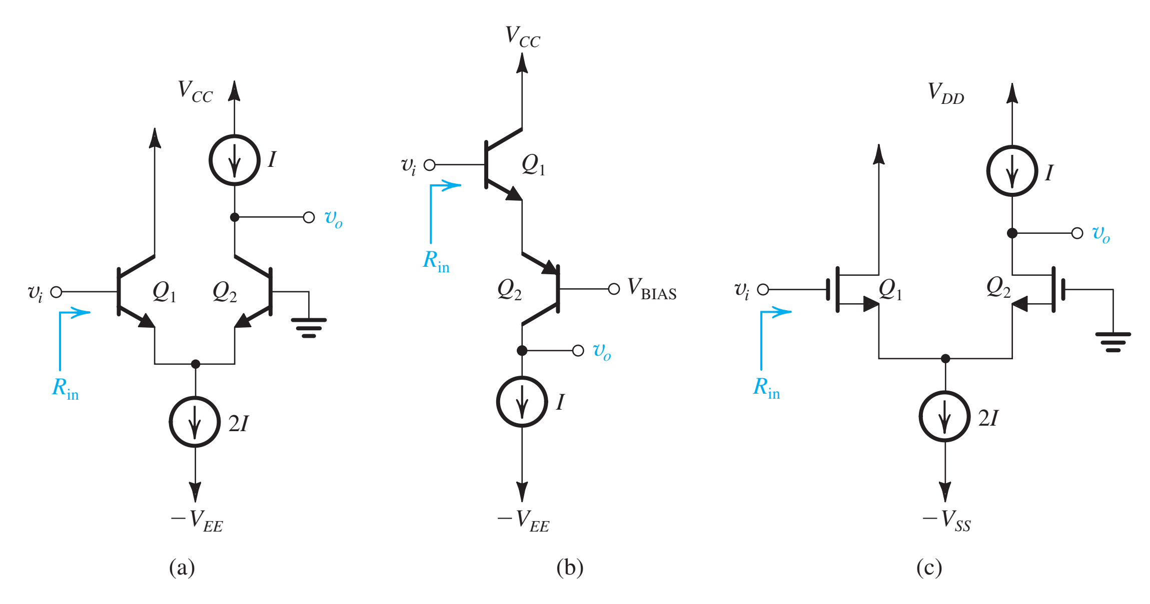


Figure 8.48 (**a**) A CC–CB amplifier. (**b**) Another version of the CC–CB circuit with Q_2 implemented using a *pnp* transistor. (**c**) The MOSFET version of the circuit in (**a**).

dc voltage at the collector of Q_2. Both issues are usually looked after in the larger circuit of which the CC–CB amplifier is a part.

An interesting version of the CC–CB configuration is shown in Fig. 8.48(b). Here the CB stage is implemented with a *pnp* transistor. Although only one current source is now needed, observe that we also need to establish an appropriate bias voltage at the base of Q_2. This circuit is part of the internal circuit of the popular 741 op amp, which will be studied in Chapter 13.

The MOSFET version of the circuit in Fig. 8.48(a) is the CD–CG amplifier shown in Fig. 8.48(c).

Example 8.8

For the CC–CB amplifiers in Fig. 8.48(a) and (b), find R_{in}, v_o/v_i, and v_o/v_{sig} when each amplifier is fed with a signal source having a resistance R_{sig}, and a load resistance R_L is connected at the output. For simplicity, neglect r_o.

Solution

The analysis of both circuits is illustrated in Fig. 8.49. Observe that both amplifiers have the same R_{in} and v_o/v_i. The overall voltage gain v_o/v_{sig} can be found as

$$\frac{v_o}{v_{sig}} = \frac{R_{in}}{R_{in}+R_{sig}} \frac{\alpha_2 R_L}{2r_e}$$

Example 8.8 *continued*

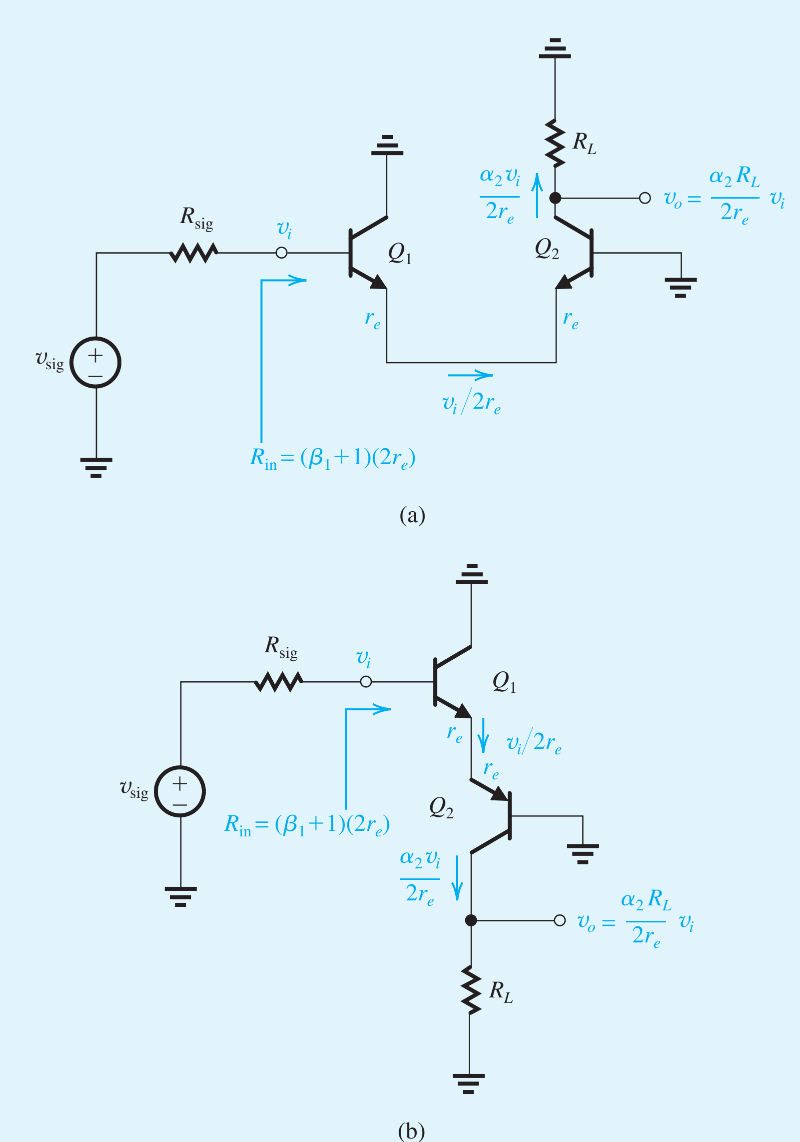


Figure 8.49 Circuits for Example 8.8.

EXERCISES

8.31 For the amplifiers in Example 8.8 find R_{in}, v_o/v_i, and v_o/v_{sig} for the case $I = 1$ mA, $\beta = 100$. $R_L = R_{sig} = 5\ \text{k}\Omega$.
Ans. 5.05 kΩ; 100 V/V; 50 V/V

D8.32 (a) Neglecting r_{o1} and the body effect, show that the voltage gain v_o/v_i of the CD–CG amplifier shown earlier in Fig. 8.48(c) is given by

$$\frac{v_o}{v_i} = \frac{IR_L}{V_{OV}}$$

where R_L is a load resistance connected at the output and V_{OV} is the overdrive voltage at which each of Q_1 and Q_2 is operating.
(b) For $I = 0.1$ mA and $R_L = 20$ kΩ, find W/L for each of Q_1 and Q_2 to obtain a gain of 10 V/V. Assume $k_n' = 200\ \mu A/V^2$.
Ans. (b) $W/L = 25$

Summary

- Integrated-circuit fabrication technology offers the circuit designer many exciting opportunities, the most important of which is the large number of inexpensive small-area MOS transistors. An overriding concern for IC designers, however, is the minimization of chip area or "silicon real estate." As a result, large-valued resistors and capacitors are virtually absent.
- Biasing in integrated circuits utilizes current sources. As well, current sources are used as load devices. Typically an accurate and stable reference current is generated and then replicated to provide bias currents for the various amplifier stages on the chip. The heart of the current-steering circuitry utilized to perform this function is the current mirror.
- The MOS current mirror has a current transfer ratio of $(W/L)_2/(W/L)_1$. For a bipolar mirror, the ratio is I_{S2}/I_{S1}.
- Bipolar mirrors suffer from the finite β, which reduces the accuracy of the current transfer ratio.
- Both bipolar and MOS mirrors of the basic type have a finite output resistance equal to r_o of the output device. Also, for proper operation, a voltage of at least 0.3 V is required across the output transistor of a simple bipolar mirror ($|V_{OV}|$ for the MOS case).
- The basic gain cell of IC amplifiers is the CS (CE) amplifier with a current-source load. For an ideal current-source load (i.e., one with infinite output resistance), the transistor operates in an open-circuit fashion and thus provides the maximum gain possible, $A_{vo} = -g_m r_o = -A_0$.
- The intrinsic gain A_0 is given by $A_0 = V_A/V_T$ for a BJT and $A_0 = V_A/(V_{OV}/2)$ for a MOSFET. For a BJT, A_0 is constant independent of bias current and device dimensions. For a MOSFET, A_0 is inversely proportional to $\sqrt{I_D}$ (see Eq. 8.46).
- Simple current-source loads reduce the gain realized in the basic gain cell because of their finite output resistance (usually comparable to the value of r_o of the amplifying transistor).
- To raise the output resistance of the CS or CE transistor, we stack a CG or CB transistor on top. This is cascoding.
- The CG and CB amplifiers act as current buffers. They have a short-circuit current gain of unity or, equivalently, a short-circuit transconductance equal to g_m of the transistor. For the CG: $R_{in} = \dfrac{r_o + R_L}{g_m r_o}$ and $R_{out} = R_s + r_o + g_m r_o R_s$. For the CB: $R_{in} = r_e \dfrac{r_o + R_L}{r_o + \dfrac{R_L}{\beta + 1}}$ and $R_{out} = (R_e \| r_\pi) + r_o + g_m r_o (R_e \| r_\pi)$.
- The CG or CB transistor in the cascode passes the current $g_{m1} v_i$ provided by the CS or CE transistor to the output but increases the resistance at the output from r_{o1} to $(g_{m2} r_{o2}) r_{o1}$ in the MOS case [$g_{m2} r_{o2} (r_{o1} \| r_{\pi 2})$ in the bipolar case]. The maximum output resistance achieved in the bipolar case is $\beta_2 r_{o2}$.
- A MOS cascode amplifier operating with an ideal current-source load achieves a gain of $(g_m r_o)^2 = A_0^2$.
- To realize the full advantage of cascoding, the load current-source must also be cascoded, in which case a gain as high as $\frac{1}{2}A_0^2$ can be obtained.
- Double cascoding is possible in the MOS case only. However, the large number of transistors in the

stack between the power-supply rails results in the disadvantage of a severely limited output-signal swing. The folded-cascode configuration helps resolve this issue.

- A CS amplifier with a resistance R_s in its source lead has an output resistance $R_o \simeq (1 + g_m R_s) r_o$. The corresponding formula for the BJT case is $R_o = [1 + g_m (R_e \| r_\pi)] r_o$.
- Cascoding can be applied to current mirrors to increase their output resistances. An alternative that also solves the β problem in the bipolar case is the Wilson circuit. The MOS Wilson mirror has an output resistance of $(g_m r_o) r_o$, and the BJT version has an output resistance of $\frac{1}{2}\beta r_o$. Both the cascode and Wilson mirrors require at least 1 V or so for proper operation.
- The Widlar current source provides an area-efficient way to implement a low-valued constant-current source that also has a high output resistance.
- Preceding the CE (CS) transistor with an emitter follower (a source follower) results in increased input resistance in the BJT case and wider bandwidth in both the BJT and MOS cases.
- Preceding the CB (CG) transistor with an emitter follower (a source follower) solves the low-input-resistance problem of the CB and CG configurations.
- The Darlington configuration results in an equivalent BJT with a current gain approaching β^2.

PROBLEMS

Computer Simulation Problems

SIM Problems identified by the multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as allowable signal swing and amplifier nonlinear distortion. Instructions to assist in setting up PSpice and Multism simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 8.2: IC Biasing—Current Sources, Current Mirrors, and Current-Steering Circuits

D 8.1 For $V_{DD} = 1.3$ V and using $I_{REF} = 100\ \mu$A, it is required to design the circuit of Fig. 8.1 to obtain an output current whose nominal value is 100 μA. Find R if Q_1 and Q_2 are matched with channel lengths of 0.5 μm, channel widths of 5 μm, $V_t = 0.4$ V, and $k_n' = 500\ \mu\text{A/V}^2$. What is the lowest possible value of V_O? Assuming that for this process technology the Early voltage $V_A' = 5$ V/μm, find the output resistance of the current source. Also, find the change in output current resulting from a +0.5-V change in V_O.

D 8.2 Using $V_{DD} = 1.8$ V and a pair of matched MOSFETs, design the current-source circuit of Fig. 8.1 to provide an output current of 150-μA nominal value. To simplify matters, assume that the nominal value of the output current is obtained at $V_O \simeq V_{GS}$. It is further required that the circuit operate for V_O in the range of 0.3 V to V_{DD} and that the change in I_O over this range be limited to 10% of the nominal value of I_O. Find the required value of R and the device dimensions. For the fabrication-process technology utilized, $\mu_n C_{ox} = 400\ \mu\text{A/V}^2$, $V_A' = 10$ V/μm, and $V_t = 0.5$ V.

D 8.3 Sketch the *p*-channel counterpart of the current-source circuit of Fig. 8.1. Note that while the circuit of Fig. 8.1 should more appropriately be called a current sink, the corresponding PMOS circuit is a current source. Let $V_{DD} = 1.3$ V, $|V_t| = 0.4$ V, Q_1 and Q_2 be matched, and $\mu_p C_{ox} = 80\ \mu\text{A/V}^2$. Find the device W/L ratios and the value of the resistor that sets the value of I_{REF} so that a nominally 80-μA output current is obtained. The current source is required to operate for V_O as high as 1.1 V. Neglect channel-length modulation.

SIM **8.4** Consider the current-mirror circuit of Fig. 8.2 with two transistors having equal channel lengths but with Q_2 having a width five times that of Q_1. If I_{REF} is 20 μA and the transistors are operating at an overdrive voltage of 0.2 V, what I_O results? What is the minimum allowable value of V_O for proper operation of the current source? If $V_t = 0.5$ V, at what value of V_O will the nominal value of I_O be obtained? If V_O increases by 1 V, what is the corresponding increase in I_O? Let $V_A = 20$ V.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

8.5 For the current-steering circuit of Fig. P8.5, find I_O in terms of I_{REF} and device W/L ratios.

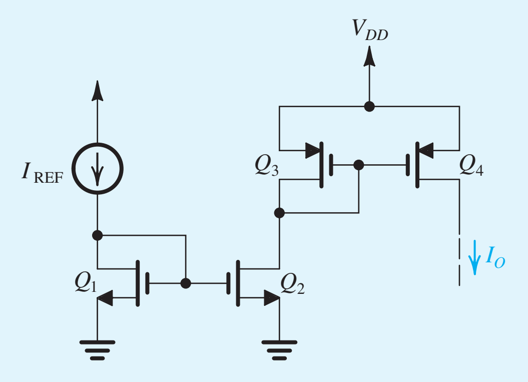


Figure P8.5

D 8.6 The current-steering circuit of Fig. P8.6 is fabricated in a CMOS technology for which $\mu_n C_{ox} = 400\ \mu A/V^2$, $\mu_p C_{ox} = 100\ \mu A/V^2$, $V_{tn} = 0.5$ V $V_{tp} = -0.5$ V, $V'_{An} = 5$ V/μ m, and $|V'_{Ap}| = 5$ V/μm. If all devices have $L = 0.5\ \mu m$, design the circuit so that $I_{REF} = 20\ \mu A$, $I_2 = 100\ \mu A$, $I_3 = I_4 = 40\ \mu A$, and $I_5 = 80\ \mu A$. Use the minimum possible device widths needed to achieve proper operation of the current source Q_2 for voltages at its drain as high as +0.8 V and proper operation of the current sink Q_5 with voltages at its drain as low as –0.8 V. Specify the widths of all devices and the value of R. Find the output resistance of the current source Q_2 and the output resistance of the current sink Q_5.

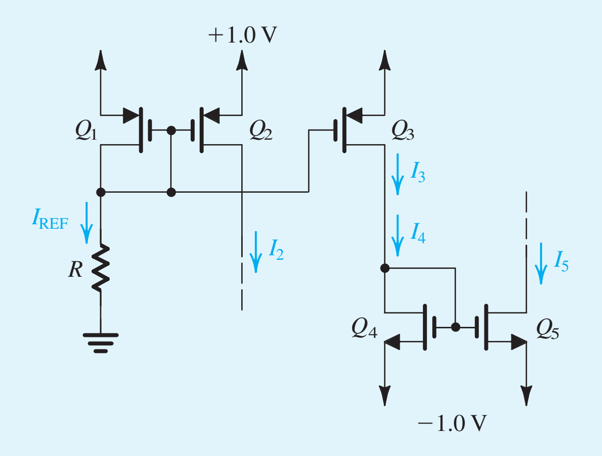


Figure P8.6

***8.7** A PMOS current mirror consists of three PMOS transistors, one diode connected and two used as current outputs. All transistors have $|V_t| = 0.6$ V, $k'_p = 100\ \mu A/V^2$, and $L = 1.0\ \mu m$ but three different widths, namely, 10 μm, 20 μm, and 40 μm. When the diode-connected transistor is supplied from a 100-μA source, how many different output currents are available? Repeat with two of the transistors diode connected and the third used to provide current output. For each possible input-diode combination, give the values of the output currents and of the V_{SG} that results.

8.8 Consider the basic bipolar current mirror of Fig. 8.7 for the case in which Q_1 and Q_2 are identical devices having $I_S = 10^{-17}$ A.

(a) Assuming the transistor β is very high, find the range of V_{BE} and I_O corresponding to I_{REF} increasing from 10 μA to 10 mA. Assume that Q_2 remains in the active mode, and neglect the Early effect.

(b) Find the range of I_O corresponding to I_{REF} in the range of 10 μA to 10 mA, taking into account the finite β. Assume that β remains constant at 100 over the current range 0.1 mA to 5 mA but that at $I_C \simeq 10$ mA and at $I_C \simeq 10\ \mu A$, $\beta = 50$. Specify I_O corresponding to $I_{REF} = 10\ \mu A$, 0.1 mA, 1 mA, and 10 mA. Note that β variation with current causes the current transfer ratio to vary with current.

8.9 Consider the basic BJT current mirror of Fig. 8.7 for the case in which Q_2 has m times the area of Q_1. Show that the current transfer ratio is given by Eq. (8.19). If β is specified to be a minimum of 80, what is the largest current transfer ratio possible if the error introduced by the finite β is limited to 10%?

8.10 Give the circuit for the *pnp* version of the basic current mirror of Fig. 8.7. If β of the *pnp* transistor is 50, what is the current gain (or transfer ratio) I_O/I_{REF} for the case of identical transistors, neglecting the Early effect?

8.11 Consider the basic BJT current mirror of Fig. 8.7 when Q_1 and Q_2 are matched and $I_{REF} = 1$ mA. Neglecting the effect of finite β, find the change in I_O, both as an absolute value and as a percentage, corresponding to V_O changing from 1 V to 10 V. The Early voltage is 90 V.

D 8.12 The current-source circuit of Fig. P8.12 utilizes a pair of matched *pnp* transistors having $I_S = 10^{-15}$ A, $\beta = 50$, and $|V_A| = 50$ V. It is required to design the circuit to provide an output current $I_O = 1$ mA at $V_O = 1$ V. What values of I_{REF} and R are needed? What is the maximum allowed value of V_O while the current source continues to operate properly? What change occurs in I_O corresponding to V_O changing from the

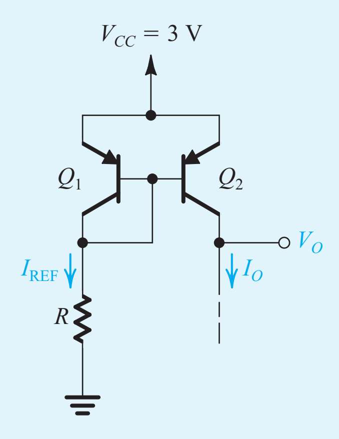


Figure P8.12

maximum positive value to –5 V? *Hint:* Adapt Eq. (8.21) for this case as:

$$I_O = I_{REF}\left[\frac{1+\dfrac{3-V_O-V_{EB}}{|V_A|}}{1+\dfrac{2}{\beta}}\right]$$

8.13 Find the voltages at all nodes and the currents through all branches in the circuit of Fig. P8.13. Assume $|V_{BE}| = 0.7$ V and $\beta = \infty$.

8.14 For the circuit in Fig. P8.14, let $|V_{BE}| = 0.7$ V and $\beta = \infty$. Find I, V_1, V_2, V_3, V_4, and V_5 for (a) $R = 10$ kΩ and (b) $R = 100$ kΩ.

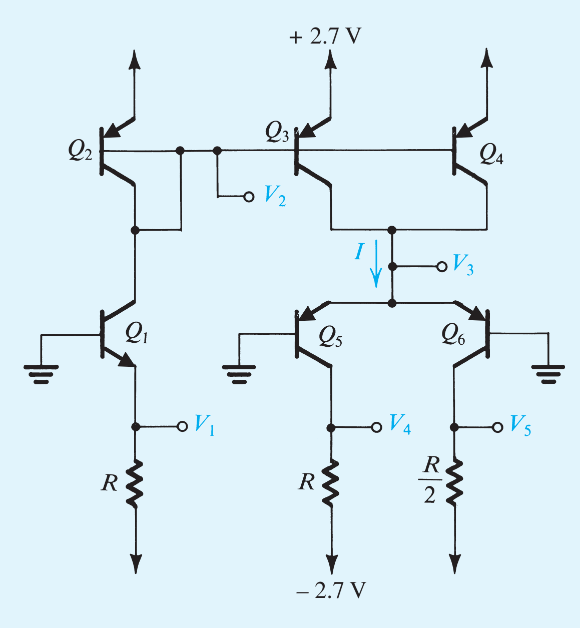


Figure P8.14

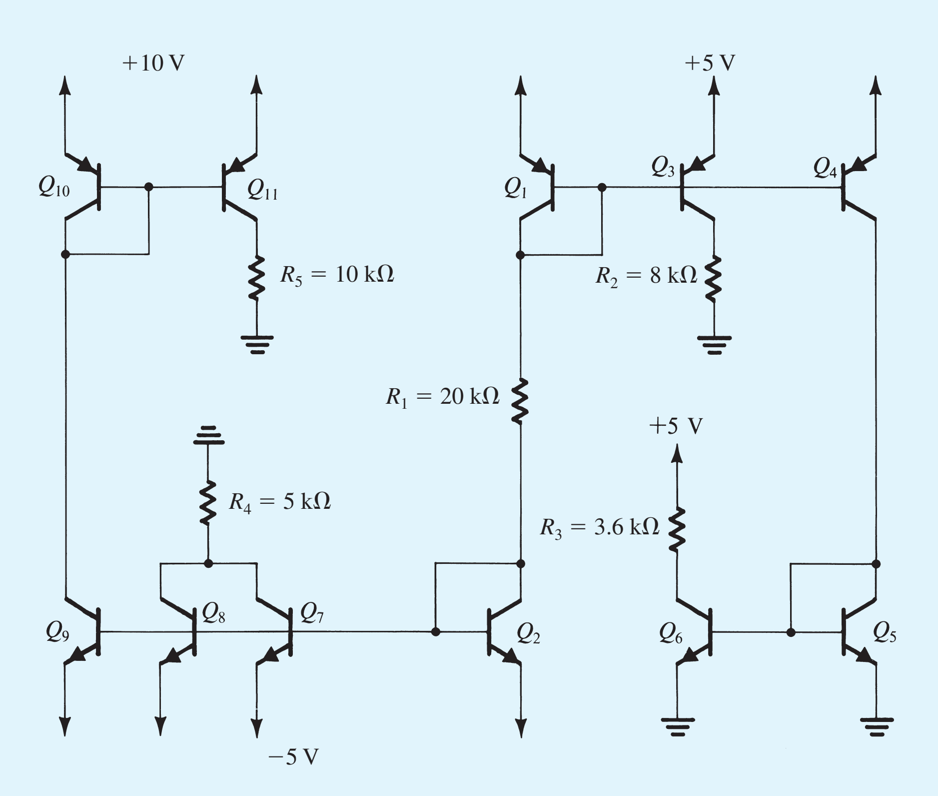


Figure P8.13

D 8.15 Using the ideas embodied in Fig. 8.10, design a multiple-mirror circuit using power supplies of ±5 V to create source currents of 0.2 mA, 0.4 mA, and 0.8 mA and sink currents of 0.5 mA, 1 mA, and 2 mA. Assume that the BJTs have $|V_{BE}| \simeq 0.7$ V and large β. What is the total power dissipated in your circuit?

***8.16** The circuit shown in Fig. P8.16 is known as a **current conveyor**.

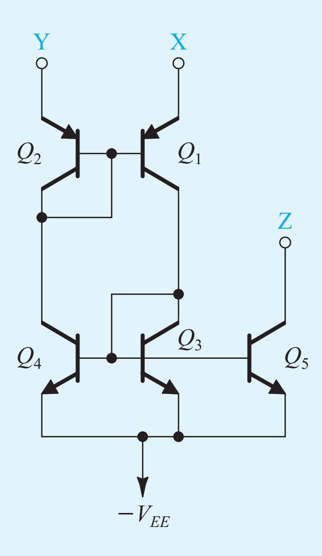


Figure P8.16

(a) Assuming that Y is connected to a voltage V, a current I is forced into X, and terminal Z is connected to a voltage that keeps Q_5 in the active region, show that a current equal to I flows through terminal Y, that a voltage equal to V appears at terminal X, and that a current equal to I flows through terminal Z. Assume β to be large; corresponding transistors are matched, and all transistors are operating in the active region.

(b) With Y connected to ground, show that a virtual ground appears at X. Now, if X is connected to a +5-V supply through a 10-kΩ resistor, what current flows through Z?

8.17 The MOSFETs in the current mirror of Fig. 8.12(a) have equal channel lengths of 0.5 μm, $W_1 = 10$ μm, $W_2 = 50$ μm, $\mu_n C_{ox} = 500$ μA/V^2, and $V_A' = 10$ V/μm. If the input bias current is 100 μA, find R_{in}, A_{is}, and R_o.

D 8.18 The MOSFETs in the current mirror of Fig. 8.12(a) have equal channel lengths, $\mu_n C_{ox} = 400$ μA/V^2 and $V_A' = 20$ V/μm. If the input bias current is 200 μA, find W_1, W_2, and L to obtain a short-circuit current gain of 4, an input resistance of 500 Ω, and an output resistance of 20 kΩ.

8.19 Figure P8.19 shows an amplifier utilizing a current mirror Q_2–Q_3. Here Q_1 is a common-source amplifier fed with $v_I = V_{GS} + v_i$, where V_{GS} is the gate-to-source dc bias voltage of Q_1 and v_i is a small signal to be amplified. Find the signal component of the output voltage v_O and hence the small-signal voltage gain v_o/v_i. Also, find the small-signal resistance of the diode-connected transistor Q_2 in terms of g_{m2}, and hence the total resistance between the drain of Q_1 and ground. What is the voltage gain of the CS amplifier Q_1? Neglect all r_o's.

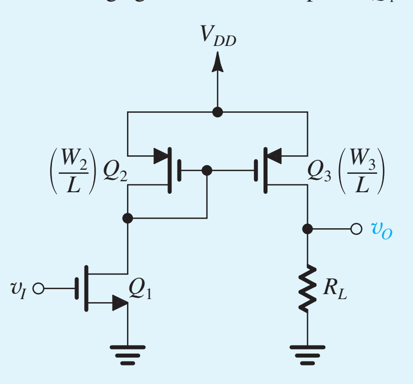


Figure P8.19

***8.20** Figure P8.20 shows a current-mirror circuit prepared for small-signal analysis. Replace the BJTs with their hybrid-π models and find expressions for R_{in}, i_o/i_i, and R_o, where i_o is the output short-circuit current. Assume $r_o \gg r_\pi$.

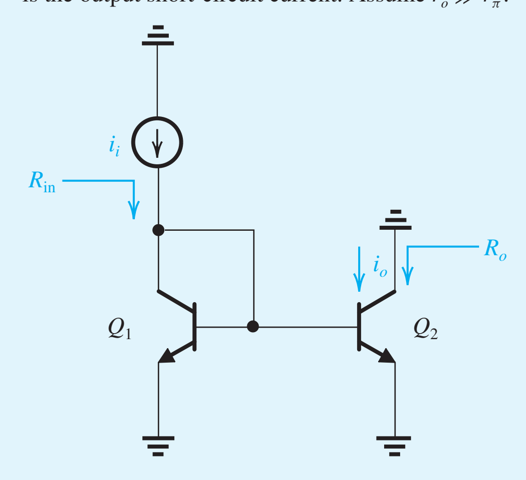


Figure P8.20

8.21 It is required to find the incremental (i.e., small-signal) resistance of each of the diode-connected transistors shown in

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

Fig. P8.21. Assume that the dc bias current $I = 0.1$ mA. For the MOSFET, let $\mu_n C_{ox} = 200\ \mu\text{A/V}^2$ and $W/L = 10$. Neglect r_o for both devices.

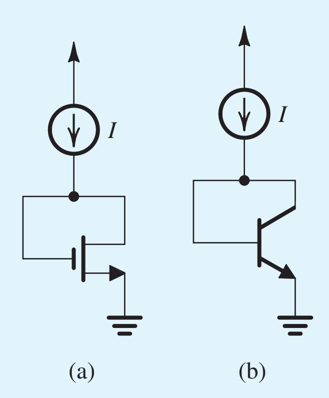


Figure P8.21

8.22 For the base-current-compensated mirror of Fig. 8.11, let the three transistors be matched and specified to have a collector current of 1 mA at $V_{BE} = 0.7$ V. For I_{REF} of 100 μA and assuming $\beta = 100$, what will the voltage at node x be? If I_{REF} is increased to 1 mA, what is the change in V_x? What is the value of I_O obtained with $V_O = V_x$ in both cases? Give the percentage difference between the actual and ideal value of I_O. What is the lowest voltage at the output for which proper current-source operation is maintained?

D 8.23 Extend the current-mirror circuit of Fig. 8.11 to n outputs. What is the resulting current transfer ratio from the input to each output, I_O/I_{REF}? If the deviation from unity is to be kept at 0.2% or less, what is the maximum possible number of outputs for BJTs with $\beta = 150$?

***8.24** For the base-current-compensated mirror of Fig. 8.11, show that the incremental input resistance (seen by the reference current source) is approximately $2V_T/I_{\text{REF}}$. Evaluate R_{in} for $I_{\text{REF}} = 100\ \mu\text{A}$. (*Hint*: Q_3 is operating at a current $I_{E3} = 2I_C/\beta$, where I_C is the operating current of each of Q_1 and Q_2. Replace each transistor with its T model and neglect r_o.)

Section 8.3: The Basic Gain Cell

8.25 Find g_m, r_π, r_o, and A_0 for the CE amplifier of Fig. 8.13(b) when operated at $I = 10$ μA, 100 μA, and 1 mA. Assume $\beta = 100$ and remains constant as I is varied, and that $V_A = 10$ V. Present your results in a table.

8.26 Consider the CE amplifiers of Fig. 8.13(b) for the case of $I = 0.5$ mA, $\beta = 100$, and $V_A = 100$ V. Find R_{in}, A_{vo}, and R_o. If it is required to raise R_{in} by a factor of 5 by changing I, what value of I is required, assuming that β remains unchanged? What are the new values of A_{vo} and R_o? If the amplifier is fed with a signal source having $R_{\text{sig}} = 5\ \text{k}\Omega$ and is connected to a load of 100-kΩ resistance, find the overall voltage gain, v_o/v_{sig}.

8.27 Find the intrinsic gain of an NMOS transistor fabricated in a process for which $k_n' = 400\ \mu\text{A/V}^2$ and $V_A' = 10\ \text{V}\ \mu\text{m}$. The transistor has a 0.5-μm channel length and is operated at $V_{OV} = 0.2$ V. If a 2-mA/V transconductance is required, what must I_D and W be?

8.28 An NMOS transistor fabricated in a certain process is found to have an intrinsic gain of 50 V/V when operated at an I_D of 100 μA. Find the intrinsic gain for $I_D = 25$ μA and $I_D = 400$ μA. For each of these currents, find the factor by which g_m changes from its value at $I_D = 100$ μA.

D 8.29 Consider an NMOS transistor fabricated in a 0.18-μm technology for which $k_n' = 400\ \mu\text{A/V}^2$ and $V_A' = 5\ \text{V/}\mu\text{m}$. It is required to obtain an intrinsic gain of 20 V/V and a g_m of 2 mA/V. Using $V_{OV} = 0.2$ V, find the required values of L, W/L, and the bias current I.

D 8.30 Sketch the circuit for a current-source-loaded CS amplifier that uses a PMOS transistor for the amplifying device. Assume the availability of a single +1.8-V dc supply. If the transistor is operated with $|V_{OV}| = 0.2$ V, what is the highest instantaneous voltage allowed at the drain?

8.31 An NMOS transistor operated with an overdrive voltage of 0.25 V is required to have a g_m equal to that of an *npn* transistor operated at $I_C = 0.1$ mA. What must I_D be? What value of g_m is realized?

8.32 For an NMOS transistor with $L = 1$ μm fabricated in the 0.5-μm process specified in Table J.1 in Appendix J, find g_m, r_o, and A_0 if the device is operated with $V_{OV} = 0.5$ V and $I_D = 100$ μA. Also, find the required device width W.

8.33 For an NMOS transistor with $L = 0.3$ μm fabricated in the 0.18-μm process specified in Table J.1 in Appendix J, find g_m, r_o, and A_0 obtained when the device is operated at $I_D = 100$ μA with $V_{OV} = 0.2$ V. Also, find W.

8.34 Fill in the table below. For the BJT, let $\beta = 100$ and $V_A = 100$ V. For the MOSFET, let $\mu_n C_{ox} = 200\ \mu$A/V^2, $W/L = 40$, and $V_A = 10$ V.

	BJT Cell		MOSFET Cell	
Bias Current	$I_C = 0.1$ mA	$I_C = 1$ mA	$I_D = 0.1$ mA	$I_D = 1$ mA
g_m (mA/V)				
r_o (kΩ)				
A_0 (V/V)				
R_{in} (kΩ)				

8.35 A CS amplifier utilizes an NMOS transistor with $L = 0.54\ \mu$m and $W/L = 8$. It was fabricated in a 0.18-μm CMOS process for which $\mu_n C_{ox} = 400\ \mu$A/V^2 and $V_A' = 5$ V/μm. What is the bias current of the transistor for which $A_0 = 18$ V/V?

8.36 A CS amplifier utilizes an NMOS transistor with $L = 0.36\ \mu$m and $W/L = 8$. It was fabricated in a 0.18-μm CMOS process for which $\mu_n C_{ox} = 400\ \mu$A/V^2 and $V_A' = 5$ V/μm. Find the values of g_m and A_0 obtained at $I_D = 25\ \mu$A, 250 μA, and 2.5 mA.

D 8.37 An NMOS transistor is fabricated in the 0.18-μm process whose parameters are given in Table J.1 in Appendix J. The device has a channel length twice the minimum and is operated at $V_{OV} = 0.25$ V and $I_D = 10\ \mu$A.

(a) What values of g_m, r_o, and A_0 are obtained?
(b) If I_D is increased to 100 μA, what do V_{OV}, g_m, r_o, and A_0 become?
(c) If the device is redesigned with a new value of W so that it operates at $V_{OV} = 0.25$ V for $I_D = 100\ \mu$A, what do g_m, r_o, and A_0 become?
(d) If the redesigned device in (c) is operated at 10 μA, find V_{OV}, g_m, r_o, and A_0.
(e) Which designs and operating conditions produce the lowest and highest values of A_0? What are these values? In each of these two cases, if W/L is held at the same value but L is made 10 times larger, what gains result?

D 8.38 Find A_0 for an NMOS transistor fabricated in a CMOS process for which $k_n' = 400\ \mu$A/V^2 and $V_A' = 6$ V/μm. The transistor has a 0.5-μm channel length and is operated with an overdrive voltage of 0.15 V. What must W be for the NMOS transistor to operate at $I_D = 100\ \mu$A? Also, find the values of g_m and r_o.

D 8.39 Using a CMOS technology for which $k_n' = 200\ \mu$A/V^2 and $V_A' = 20$ V/μm, design a current-source-loaded CS amplifier for operation at $I = 50\ \mu$A with $V_{OV} = 0.2$ V. The amplifier is to have an open-circuit voltage gain of -100 V/V. Assume that the current-source load is ideal. Specify L and W/L.

D 8.40 The circuit in Fig. 8.15(a) is fabricated in a process for which $\mu_n C_{ox} = 2\mu_p C_{ox} = 200\ \mu$A/V^2, $V_{An}' = |V_{Ap}'| = 20$ V/μm, $V_{tn} = -V_{tp} = 0.5$ V, and $V_{DD} = 2.5$ V. The two transistors have $L = 0.5\ \mu$m and are to be operated at $I_D = 100\ \mu$A and $|V_{OV}| = 0.3$ V. Find the required values of V_G, $(W/L)_1$, $(W/L)_2$, and A_v.

D 8.41 The circuit in Fig. 8.15(a) is fabricated in a 0.18-μm CMOS technology for which $\mu_n C_{ox} = 400\ \mu$A/V^2, $\mu_p C_{ox} = 100\ \mu$A/V^2, $V_{tn} = -V_{tp} = 0.5$ V, $V_{An}' = 5$ V/μm, $|V_{Ap}'| = 5$ V/μm, and $V_{DD} = 1.8$ V. It is required to design the circuit to obtain a voltage gain $A_v = -40$ V/V. Use devices of equal length L operating at $I = 100\ \mu$A and $|V_{OV}| = 0.25$ V. Determine the required values of V_G, L, $(W/L)_1$, and $(W/L)_2$.

8.42 Figure P8.42 shows an IC MOS amplifier formed by cascading two common-source stages. Assuming that $V_{An} = |V_{Ap}|$ and that the biasing current sources have output resistances equal to those of Q_1 and Q_2, find an expression for the overall voltage gain in terms of g_m and r_o of Q_1 and Q_2. If Q_1 and Q_2 are to be operated at equal overdrive voltages, $|V_{OV}|$, find the required value of $|V_{OV}|$ if $|V_A| = 5$ V and the gain required is 400 V/V.

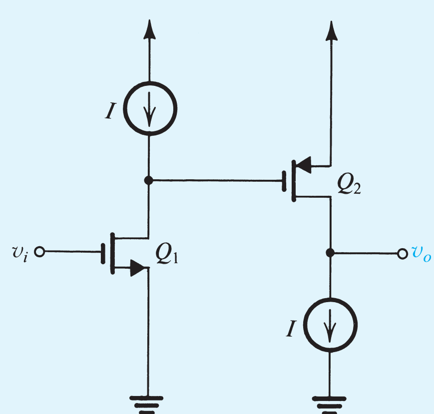


Figure P8.42

***8.43** The NMOS transistor in the circuit of Fig. P8.43 has $V_t = 0.5$ V, $k_n' W/L = 2$ mA/V^2, and $V_A = 20$ V.

(a) Neglecting the dc current in the feedback network and the effect of r_o, find V_{GS}. Then find the dc current in the feedback network and V_{DS}. Verify that you were justified in neglecting the current in the feedback network when you found V_{GS}.

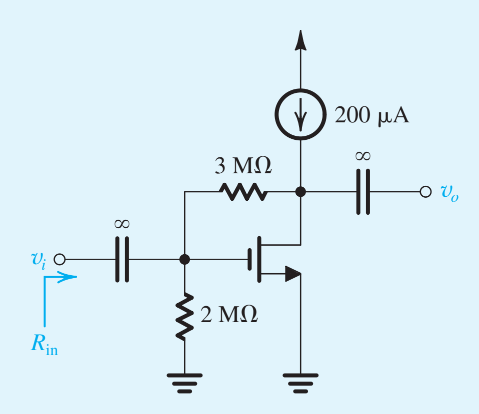


Figure P8.43

(b) Find the small-signal voltage gain, v_o/v_i. What is the peak of the largest output sine-wave signal that is possible while the NMOS transistor remains in saturation? What is the corresponding input signal?

(c) Find the small-signal input resistance R_{in}.

D 8.44 Consider the CMOS amplifier of Fig. 8.16(a) when fabricated with a process for which $k_n' = 4k_p' = 400\ \mu A/V^2$, $|V_t| = 0.5$ V, and $|V_A| = 5$ V. Find I_{REF} and $(W/L)_1$ to obtain a voltage gain of –40 V/V and an output resistance of 100 kΩ. Recall that Q_2 and Q_3 are matched. If Q_2 and Q_3 are to be operated at the same overdrive voltage as Q_1, what must their W/L ratios be?

8.45 Consider the CMOS amplifier analyzed in Example 8.4. If v_I consists of a dc bias component on which is superimposed a sinusoidal signal, find the value of the dc component that will result in the maximum possible signal swing at the output with almost-linear operation. What is the amplitude of the output sinusoid resulting? (*Note:* In practice, the amplifier would have a feedback circuit that caused it to operate at a point near the middle of its linear region.)

8.46 The power supply of the CMOS amplifier analyzed in Example 8.4 is increased to 5 V. What will the extent of the linear region at the output become?

****8.47** Consider the circuit shown in Fig. 8.16(a), using a 3.3-V supply and transistors for which $|V_t| = 0.8$ V and $L = 1\ \mu$m. For Q_1, $k_n' = 100\ \mu A/V^2$, $V_A = 100$ V, and $W = 20\ \mu$m. For Q_2 and Q_3, $k_p' = 50\ \mu A/V^2$ and $|V_A| = 50$ V. For Q_2, $W = 40\ \mu$m. For Q_3, $W = 10\ \mu$m.

(a) If Q_1 is to be biased at 100 μA, find I_{REF}. For simplicity, ignore the effect of V_A.

(b) What are the extreme values of v_O for which Q_1 and Q_2 just remain in saturation?

(c) What is the large-signal voltage gain?

(d) Find the slope of the transfer characteristic at $v_O = V_{DD}/2$.

(e) For operation as a small-signal amplifier around a bias point at $v_O = V_{DD}/2$, find the small-signal voltage gain and output resistance.

****8.48** The MOSFETs in the circuit of Fig. P8.48 are matched, having $k_n'(W/L)_1 = k_p'(W/L)_2 = 1\ mA/V^2$ and $|V_t| = 0.5$ V. The resistance $R = 1$ MΩ.

(a) For G and D open, what are the drain currents I_{D1} and I_{D2}?

(b) For $r_o = \infty$, what is the voltage gain of the amplifier from G to D? (*Hint*: Replace the transistors with their small-signal models.)

(c) For finite r_o ($|V_A| = 20$ V), what is the voltage gain from G to D and the input resistance at G?

(d) If G is driven (through a large coupling capacitor) from a source v_{sig} having a resistance of 20 kΩ, find the voltage gain v_d/v_{sig}.

(e) For what range of output signals do Q_1 and Q_2 remain in the saturation region?

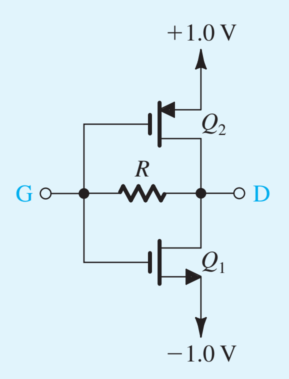


Figure P8.48

8.49 Transistor Q_1 in the circuit of Fig. P8.49 is operating as a CE amplifier with an active load provided by transistor Q_2, which is the output transistor in a current mirror formed by Q_2 and Q_3. (Note that the biasing arrangement for Q_1 is *not* shown.)

(a) Neglecting the finite base currents of Q_2 and Q_3 and assuming that their $V_{BE} \simeq 0.7$ V and that Q_2 has five times the area of Q_3, find the value of I.

(b) If Q_1 and Q_2 are specified to have $|V_A| = 30$ V, find r_{o1} and r_{o2} and hence the total resistance at the collector of Q_1.

(c) Find $r_{\pi 1}$ and g_{m1} assuming that $\beta_1 = 50$.
(d) Find R_{in}, A_v, and R_o.

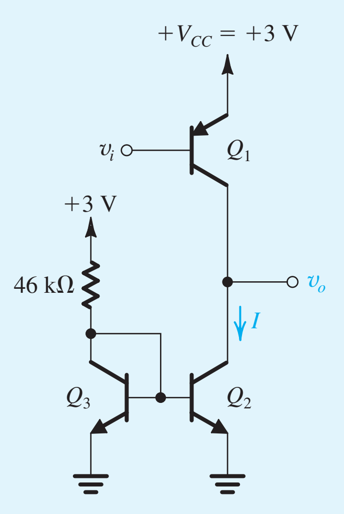


Figure P8.49

D 8.50 It is required to design the CMOS amplifier of Fig. 8.16(a) utilizing a 0.18-μm process for which $k_n' = 387\ \mu A/V^2$, $k_p' = 86\ \mu A/V^2$, $V_{tn} = -V_{tp} = 0.5$ V, $V_{DD} = 1.8$ V, $V_{An}' = 5$ V/μm, and $V_{Ap}' = -6$ V/μm. The output voltage must be able to swing to within approximately 0.2 V of the power-supply rails (i.e., from 0.2 V to 1.6 V), and the voltage gain must be at least 10 V/V. Design for a dc bias current of 50 μA, and use devices with the same channel length. If the channel length is an integer multiple of the minimum 0.18 μm, what channel length is needed and what W/L ratios are required? If it is required to raise the gain by a factor of 2, what channel length would be required, and by what factor does the total gate area of the circuit increase?

Section 8.4: The CG and CB Amplifiers

8.51 A CG amplifier operating with $g_m = 2$ mA/V and $r_o = 20$ kΩ is fed with a signal source having $R_s = 1$ kΩ and is loaded in a resistance $R_L = 20$ kΩ. Find R_{in}, R_{out}, and v_o/v_{sig}.

8.52 A CG amplifier operating with $g_m = 2$ mA/V and $r_o = 20$ kΩ is fed with a signal source having a Norton equivalent composed of a current signal i_{sig} and a source resistance $R_s = 20$ kΩ. The amplifier is loaded in a resistance $R_L = 20$ kΩ. Find R_{in} and i_o/i_{sig}, where i_o is the current through the load R_L. If R_L increases by a factor of 10, by what percentage does the current gain change? Can you see the effectiveness of the CG as a current buffer?

D 8.53 It is required to design the current source in Fig. P8.53 to deliver a current of 0.2 mA with an output resistance of 500 kΩ. The transistor has $V_A = 20$ V and $V_t = 0.5$ V. Design for $V_{OV} = 0.2$ V and specify R_s and V_{BIAS}.

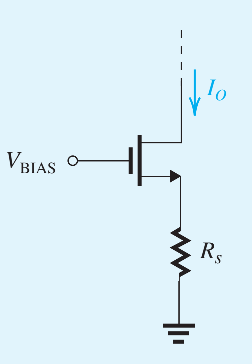


Figure P8.53

D 8.54 Figure P8.54 shows a current source realized using a current mirror with two matched transistors Q_1 and Q_2. Two equal resistances R_s are inserted in the source leads to increase the output resistance of the current source. If Q_2 is operating at $g_m = 1$ mA/V and has $V_A = 10$ V, and if the maximum allowed dc voltage drop across R_s is 0.3 V, what is the maximum available output resistance of the current source? Assume that the voltage at the common-gate node is approximately constant.

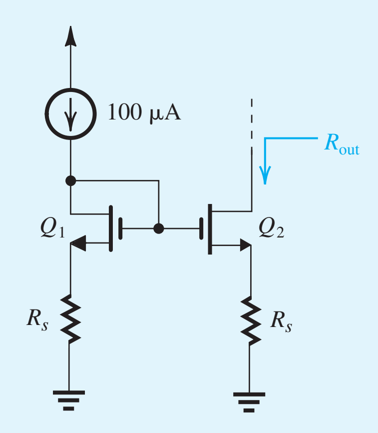


Figure P8.54

8.55 In the common-gate amplifier circuit of Fig. P8.55, Q_2 and Q_3 are matched. $k_n'(W/L)_n = k_p'(W/L)_p = 4$ mA/V^2, and all transistors have $|V_t| = 0.8$ V and $|V_A| = 20$ V.

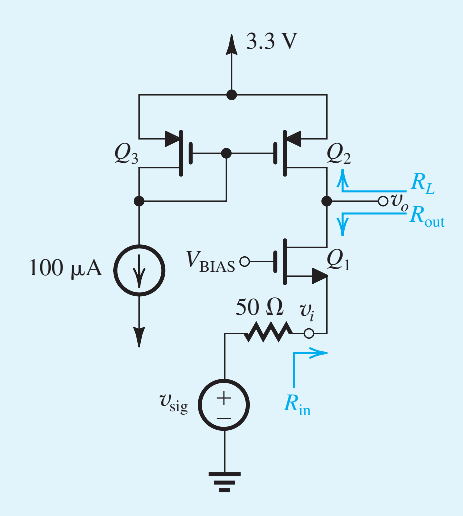


Figure P8.55

The signal v_{sig} is a small sinusoidal signal with no dc component.

(a) Neglecting the effect of V_A, find the dc drain current of Q_1 and the required value of V_{BIAS}.
(b) Find the values of g_{m1} and r_o for all transistors.
(c) Find the value of R_{in}.
(d) Find the value of R_{out}.
(e) Calculate the voltage gains v_o/v_i and v_o/v_{sig}.
(f) How large can v_{sig} be (peak-to-peak) while maintaining saturation-mode operation for Q_1 and Q_2?

8.56 For the CB amplifier, use Eq. (8.63) to explore the variation of the input resistance R_{in} with the load resistance R_L. Specifically, find R_{in} as a multiple of r_e for $R_L/r_o = 0$, 1, 10, 100, 1000, and ∞. Let $\beta = 100$. Present your results in tabular form.

8.57 What value of load resistance R_L causes the input resistance of the CB amplifier to be approximately double the value of r_e?

8.58 Show that for the CB amplifier,

$$\frac{R_{out}}{r_o} \simeq 1 + \frac{\beta(R_e/r_e)}{\beta + 1 + (R_e/r_e)}$$

Generate a table for R_{out} as a multiple of r_o versus R_e as a multiple of r_e with entries for $R_e = 0$, r_e, $2r_e$, $10r_e$, $(\beta/2)r_e$, βr_e, and $1000r_e$. Let $\beta = 100$.

8.59 As mentioned in the text, the CB amplifier functions as a current buffer. That is, when fed with a current signal, it passes it to the collector and supplies the output collector current at a high output resistance. Figure P8.59 shows a CB amplifier fed with a signal current i_{sig} having a source resistance $R_{sig} = 10\text{ k}\Omega$. The BJT is specified to have $\beta = 100$ and $V_A = 50$ V. (Note that the bias arrangement is not shown.) The output at the collector is represented by its Norton equivalent circuit. Find the value of the current gain k and the output resistance R_{out}. Note that k is the short-circuit current gain and should be evaluated using the T model of the transistor with the collector short-circuited to ground.

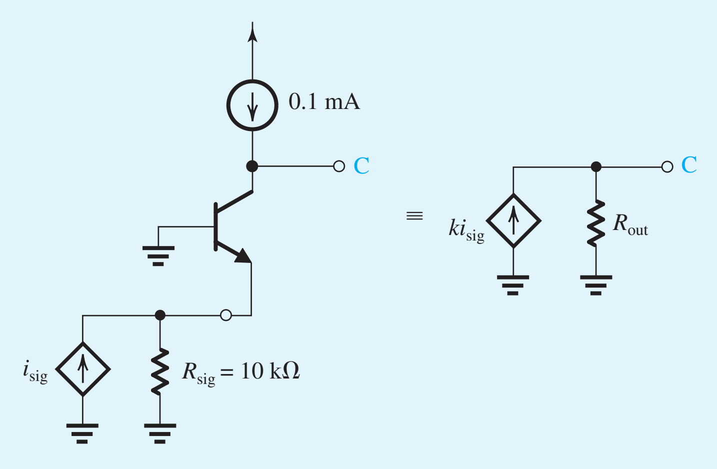


Figure P8.59

8.60 For the constant-current source circuit shown in Fig. P8.60, find the collector current I and the output resistance. The BJT is specified to have $\beta = 100$, $V_{BE} = 0.7$ V, and $V_A = 100$ V. If the collector voltage undergoes a change of 10 V while the BJT remains in the active mode, what is the corresponding change in collector current?

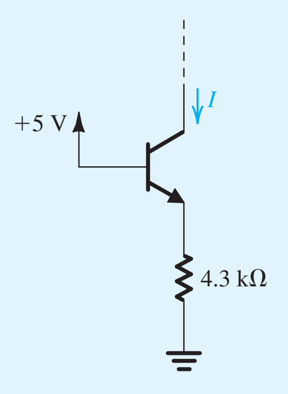


Figure P8.60

8.61 Find the value of the resistance R_e, which, when connected in the emitter lead of a CE BJT amplifier, raises the output resistance by a factor of (a) 5, (b) 10, and (c) 50. What is the maximum possible factor by which the output resistance can be raised, and at what value of R_e is it achieved? Assume the BJT has $\beta = 100$ and is biased at $I_C = 0.5$ mA.

Section 8.5: The Cascode Amplifier

D 8.62 In a MOS cascode amplifier, the cascode transistor is required to raise the output resistance by a factor of 50. If the transistor is operated at $V_{OV} = 0.2$ V, what must its V_A be? If the process technology specifies V_A' as 5 V/μm, what channel length must the transistor have?

D 8.63 For a cascode current source such as that in Fig. 8.32, show that if the two transistors are identical, the current I supplied by the current source and the output resistance R_o are related by $IR_o = 2|V_A|^2/|V_{OV}|$. Now consider the case of transistors that have $|V_A| = 4$ V and are operated at $|V_{OV}|$ of 0.2 V. Also, let $\mu_p C_{ox} = 100$ μA/V^2. Find the W/L ratios required and the output resistance realized for the two cases: (a) $I = 0.1$ mA and (b) $I = 0.5$ mA. Assume that V_{SD} for the two devices is the minimum required (i.e., $|V_{OV}|$).

D *8.64 For a cascode current source, such as that in Fig. 8.32, show that if the two transistors are identical, the current I supplied by the current source and the output resistance R_o are related by

$$IR_o = \frac{2|V_A'|^2}{|V_{OV}|} L^2$$

Now consider the case of a 0.18-μm technology for which $|V_A'| = 5$ V/μm and let the transistors be operated at $|V_{OV}| = 0.2$ V. Find the figure-of-merit IR_o for the three cases of L equal to the minimum channel length, twice the minimum, and three times the minimum. Complete the entries of the table at the bottom of the page. Give W/L and the area $2WL$ in terms of n, where n is the value of W/L for the case $I = 0.01$ mA. In the table, A_v denotes the gain obtained in a cascode amplifier such as that in Fig. 8.33 that utilizes our current source as load and which has the same values of g_m and R_o as the current-source transistors.

(a) For each current value, what is price paid for the increase in R_o and A_v obtained as L is increased?
(b) For each value of L, what advantage is obtained as I is increased, and what is the price paid? (*Hint:* We will see in Chapter 10 that the amplifier bandwidth increases with g_m.)
(c) Contrast the performance obtained from the circuit with the largest area with that obtained from the circuit with the smallest area.

D 8.65 Design the cascode amplifier of Fig. 8.30(a) to obtain $g_{m1} = 2$ mA/V and $R_o = 200$ kΩ. Use a 0.18-μm technology for which $V_{tn} = 0.5$ V, $V_A' = 5$ V/μm, and $k_n' = 400$ μA/V^2. Determine L, W/L, V_{G2}, and I. Use identical transistors operated at $V_{OV} = 0.25$ V, and design for the maximum possible negative signal swing at the output. What is the value of the minimum permitted output voltage?

8.66 The cascode amplifier of Fig. 8.33 is operated at a current of 0.2 mA with all devices operating at $|V_{OV}| = 0.20$ V.

	$L = L_{min} = 0.18$ μm $IR_o =$ V				$L = 2L_{min} = 0.36$ μm $IR_o =$ V				$L = 3L_{min} = 0.54$ μm $IR_o =$ V			
	g_m (mA/V)	R_o (kΩ)	A_v (V/V)	$2WL$ (μm^2)	g_m (mA/V)	R_o (kΩ)	A_v (V/V)	$2WL$ (μm^2)	g_m (mA/V)	R_o (kΩ)	A_v (V/V)	$2WL$ (μm^2)
$I = 0.01$ mA $W/L = n$												
$I = 0.1$ mA $W/L =$												
$I = 1.0$ mA $W/L =$												

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

All devices have $|V_A| = 4$ V. Find g_{m1}, the output resistance of the amplifier, R_{on}, the output resistance of the current source, R_{op}, the overall output resistance, R_o, and the voltage gain, A_v.

D 8.67 Design the CMOS cascode amplifier in Fig. 8.33 for the following specifications: $g_{m1} = 1$ mA/V and $A_v = -280$ V/V. Assume that for the available fabrication process, $|V_A'| = 5$ V/μm for both NMOS and PMOS devices and that $\mu_n C_{ox} = 4\,\mu_p C_{ox} = 400\ \mu\text{A/V}^2$. Use the same channel length L for all devices and operate all four devices at $|V_{OV}| = 0.25$ V. Determine the required channel length L, the bias current I, and the W/L ratio for each of four transistors. Assume that suitable bias voltages have been chosen, and neglect the Early effect in determining the W/L ratios.

D 8.68 Design the circuit of Fig. 8.32 to provide an output current of 100 μA. Use $V_{DD} = 3.3$ V, and assume the PMOS transistors to have $\mu_p C_{ox} = 60\ \mu\text{A/V}^2$, $V_{tp} = -0.8$ V, and $|V_A| = 5$ V. The current source is to have the widest possible signal swing at its output. Design for $V_{OV} = 0.2$ V, and specify the values of the transistor W/L ratios and of V_{G3} and V_{G4}. What is the highest allowable voltage at the output? What is the value of R_o?

8.69 The cascode transistor can be thought of as providing a "shield" for the input transistor from the voltage variations at the output. To quantify this "shielding" property of the cascode, consider the situation in Fig. P8.69. Here we have grounded the input terminal (i.e., reduced v_i to zero), applied a small change v_x to the output node, and denoted the voltage change that results at the drain of Q_1 by v_y. By what factor is v_y smaller than v_x?

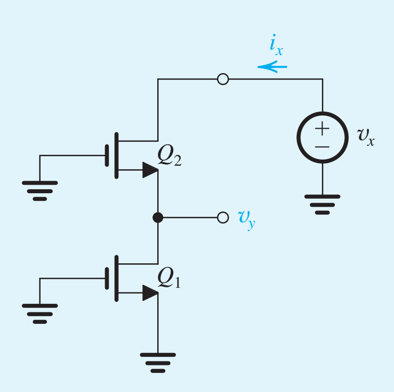


Figure P8.69

***8.70** In this problem we investigate whether, as an alternative to cascoding, we can simply increase the channel length L of the CS MOSFET. Specifically, we wish to compare the two circuits shown in Fig. P8.70(b) and (c). The circuit in Fig. P8.70(b) is a CS amplifier in which the channel length has been quadrupled relative to that of the original CS amplifier in Fig. P8.70(a) while the drain bias current has been kept constant.

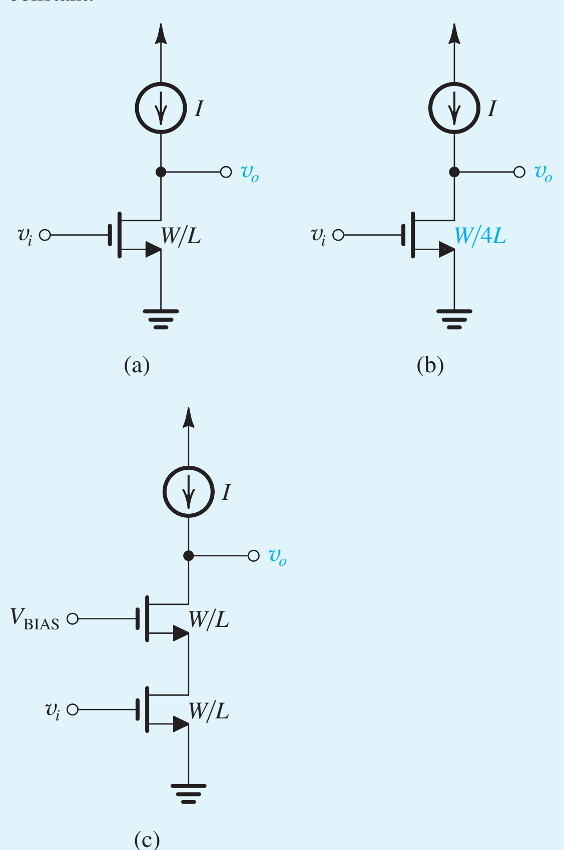


Figure P8.70

(a) Show that for this circuit V_{OV} is double that of the original circuit, g_m is half that of the original circuit, and $\frac{v_o}{v_i}$ is double that of the original circuit.

(b) Compare these values to those of the cascode circuit in Fig. P8.70(c), which is operating at the same bias current and has the same minimum voltage requirement at the drain as in the circuit of Fig. P8.70(b).

8.71 Consider the cascode amplifier of Fig. 8.33 with the dc component at the input $V_I = 0.6$ V, $V_{G2} = 0.9$ V, $V_{G3} = 0.4$ V, $V_{G4} = 0.7$ V, and $V_{DD} = 1.3$ V. If all devices are matched, that is, $k_{n1} = k_{n2} = k_{p3} = k_{p4}$, and have equal $|V_t|$ of 0.4 V, what is the overdrive voltage at which the four transistors are operating? What is the allowable voltage range at the output?

SIM **8.72** A CMOS cascode amplifier such as that in Fig. 8.34(a) has identical CS and CG transistors that have $W/L = 5.4\ \mu m/0.36\ \mu m$ and biased at $I = 0.2$ mA. The fabrication process has $\mu_n C_{ox} = 400\ \mu A/V^2$, and $V'_A = 5\ V/\mu m$. At what value of R_L does the gain become −100 V/V? What is the voltage gain of the common-source stage?

8.73 The purpose of this problem is to investigate the signal currents and voltages at various points throughout a cascode amplifier circuit. Knowledge of this signal distribution is very useful in designing the circuit so as to allow for the required signal swings. Figure P8.73 shows a CMOS cascode amplifier with all dc voltages replaced with signal grounds. As well, we have explicitly shown the resistance r_o of each of the four transistors. For simplicity, we are assuming that the four transistors have the same g_m and r_o. The amplifier is fed with a signal v_i.

(a) Determine R_1, R_2, and R_3. Assume $g_m r_o \gg 1$.
(b) Determine i_1, i_2, i_3, i_4, i_5, i_6, and i_7, all in terms of v_i. (*Hint:* Use the current-divider rule at the drain of Q_1.)
(c) Determine v_1, v_2, and v_3, all in terms of v_i.
(d) If v_i is a 5-mV peak sine wave and $g_m r_o = 20$, sketch and clearly label the waveforms of v_1, v_2, and v_3.

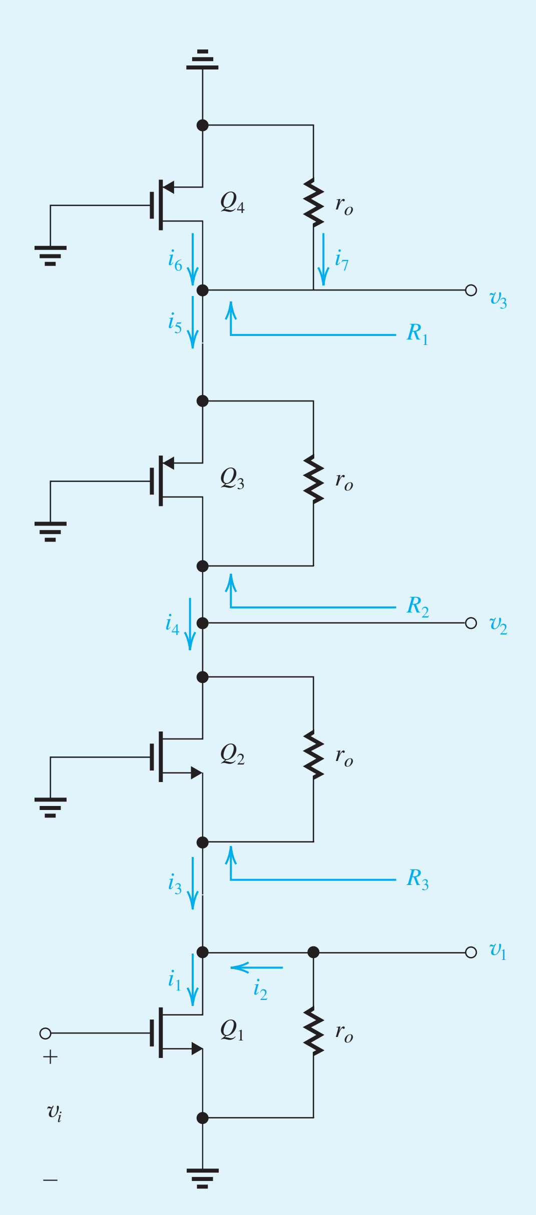


Figure P8.73

D 8.74 Design the double-cascode current source shown in Fig. P8.74 to provide $I = 0.2$ mA and the largest possible signal swing at the output; that is, design for the minimum allowable voltage across each transistor. The 0.13-μm CMOS fabrication process available has $V_{tp} = -0.4$ V, $V'_A = -6\ V/\mu m$, and $\mu_p C_{ox} = 100\ \mu A/V^2$. Use devices with $L = 0.4\ \mu m$, and operate at $|V_{OV}| = 0.2$ V. Specify V_{G1}, V_{G2}, V_{G3}, and the W/L ratios of the transistors. What is the value of R_o achieved?

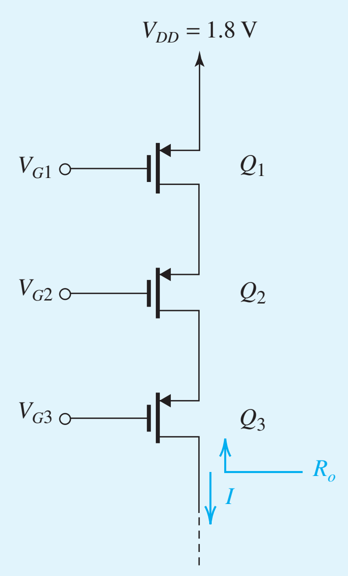


Figure P8.74

***8.75** Figure P8.75 shows a folded-cascode CMOS amplifier utilizing a simple current source Q_2, supplying a current $2I$, and a cascoded current source (Q_4, Q_5) supplying a current I. Assume, for simplicity, that all transistors have equal parameters g_m and r_o.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

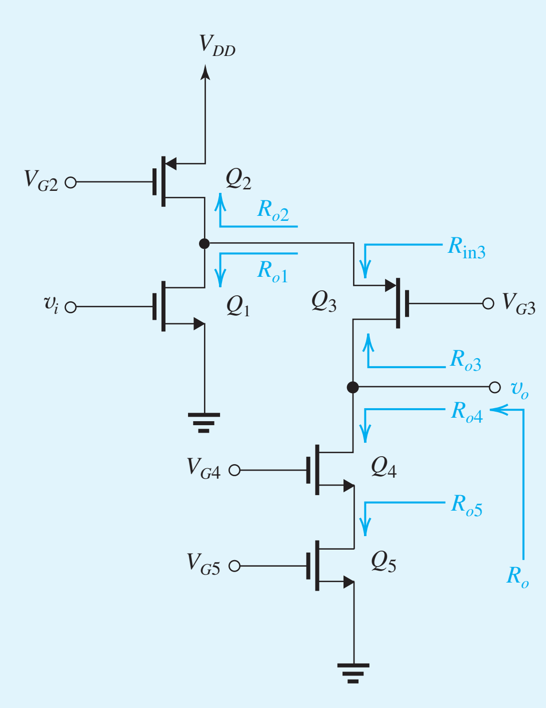


Figure P8.75

(a) Give approximate expressions for all the resistances indicated.

(b) Find the amplifier output resistance R_o.

(c) Show that the short-circuit transconductance G_m is approximately equal to g_{m1}. Note that the short-circuit transconductance is determined by short-circuiting v_o to ground and finding the current that flows through the short circuit, $G_m v_i$.

(d) Find the overall voltage gain v_o/v_i and evaluate its value for the case $g_{m1} = 2$ mA/V and $A_0 = 30$.

8.76 A cascode current source formed of two *pnp* transistors for which $\beta = 50$ and $V_A = 5$ V supplies a current of 0.2 mA. What is the output resistance?

8.77 Use Eq. (8.88) to show that for a BJT cascode current source utilizing identical *pnp* transistors and supplying a current I,

$$IR_o = \frac{|V_A|}{(V_T/|V_A|) + (1/\beta)}$$

Evaluate the figure-of-merit IR_o for the case $|V_A| = 5$ V and $\beta = 50$. Now find R_o for the cases of $I = 0.1$, 0.5, and 1.0 mA.

8.78 Consider the BJT cascode amplifier of Fig. 8.38 for the case all transistors have equal β and r_o. Show that the voltage gain A_v can be expressed in the form

$$A_v = -\frac{1}{2}\frac{|V_A|/V_T}{(V_T/|V_A|) + (1/\beta)}$$

Evaluate A_v for the case $|V_A| = 5$ V and $\beta = 50$. Note that except for the fact that β depends on I as a second-order effect, the gain is independent of the bias current I!

8.79 A bipolar cascode amplifier has a current-source load with an output resistance βr_o. Let $\beta = 50$, $|V_A| = 100$ V, and $I = 0.2$ mA. Find the voltage gain A_v.

D *8.80 Figure P8.80 shows four possible realizations of the folded cascode amplifier. Assume that the BJTs have $\beta = 100$ and that both the BJTs and the MOSFETs have $|V_A| = 5$ V.

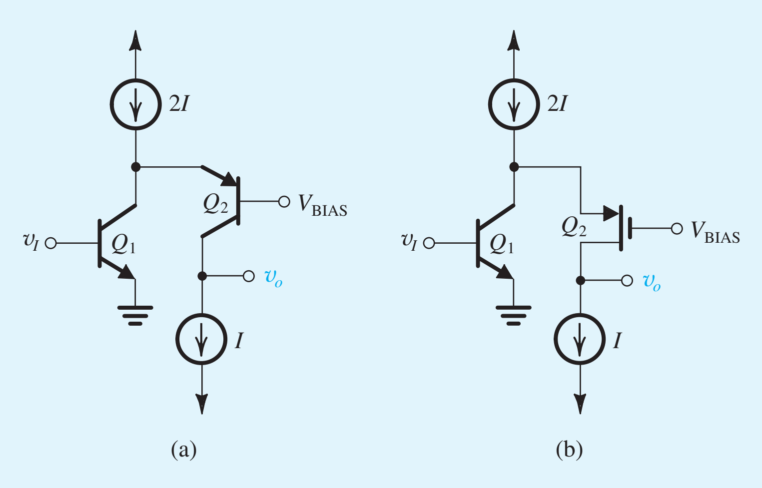


Figure P8.80

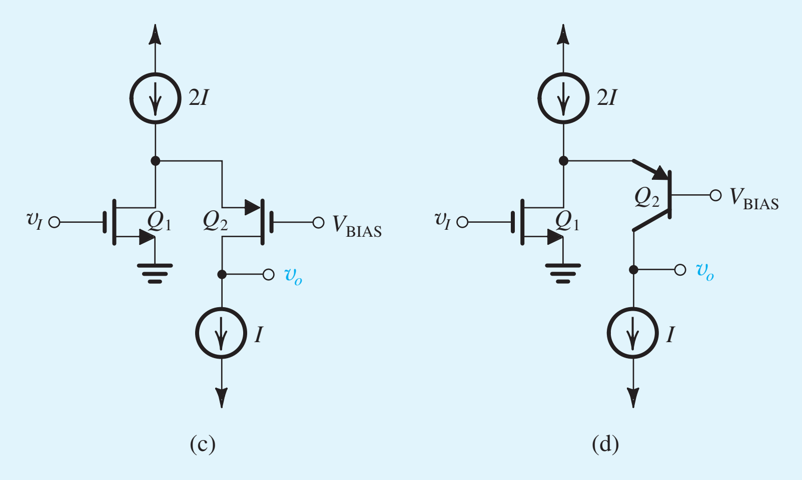


Figure P8.80 *continued*

Let $I = 100\ \mu$A, and assume that the MOSFETs are operating at $|V_{OV}| = 0.2$ V. Assume the current sources are ideal. For each circuit determine R_{in}, R_o, and A_{vo}. Comment on your results.

8.81 In this problem, we will explore the difference between using a BJT as cascode device and a MOSFET as cascode device. Refer to Fig. P8.81. Given the following data, calculate G_m, R_o, and A_{vo} for the circuits (a) and (b):

$I = 100\ \mu$A, $\beta = 125$, $\mu_n C_{ox} = 400\ \mu$A/V^2, $W/L = 25$, $V_A = 1.8$ V

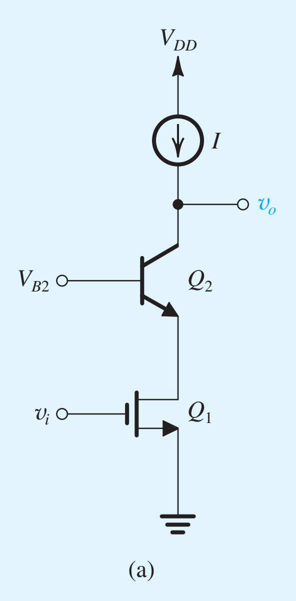


Figure P8.81

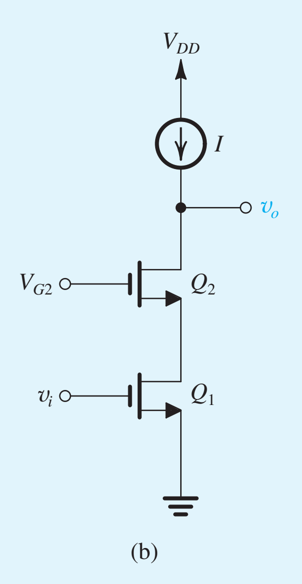


Figure P8.81 *continued*

Section 8.6: Current-Mirror Circuits with Improved Performance

SIM **8.82** In a particular cascoded current mirror, such as that shown in Fig. 8.39, all transistors have $V_t = 0.6$ V, $\mu_n C_{ox} = 160\ \mu$A/V^2, $L = 1\ \mu$m, and $V_A = 10$ V. Width $W_1 = W_4 = 4\ \mu$m, and $W_2 = W_3 = 40\ \mu$m. The reference current I_{REF} is $20\ \mu$A. What output current results? What are the voltages at the gates of Q_2 and Q_3? What is the lowest voltage at the output for which current-source operation is possible? What are the values of g_m and

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

r_o of Q_2 and Q_3? What is the output resistance of the mirror?

8.83 Find the output resistance of the double-cascode current mirror of Fig. P8.83.

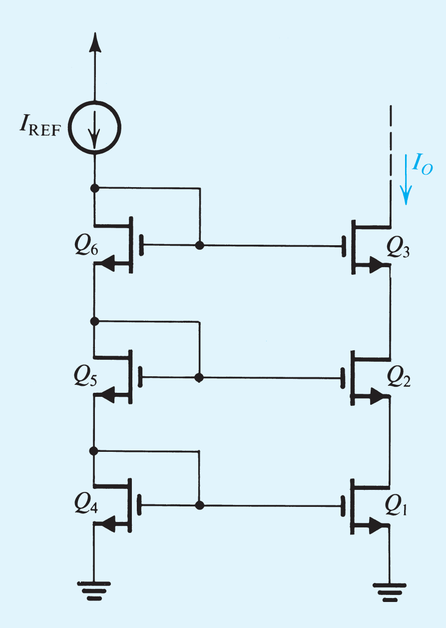


Figure P8.83

8.84 Consider the Wilson current-mirror circuit of Fig. 8.40 when supplied with a reference current I_{REF} of 1 mA. What is the change in I_O corresponding to a change of +10 V in the voltage at the collector of Q_3? Give both the absolute value and the percentage change. Let $\beta = 100$ and $V_A = 100$ V.

D 8.85 (a) The circuit in Fig. P8.85 is a modified version of the Wilson current mirror. Here the output transistor is "split" into two matched transistors, Q_3 and Q_4. Find I_{O1} and I_{O2} in terms of I_{REF}. Assume all transistors to be matched with current gain β.
(b) Use this idea to design a circuit that generates currents of 0.1 mA, 0.2 mA, and 0.4 mA, using a reference current source of 0.7 mA. What are the actual values of the currents generated for $\beta = 50$?

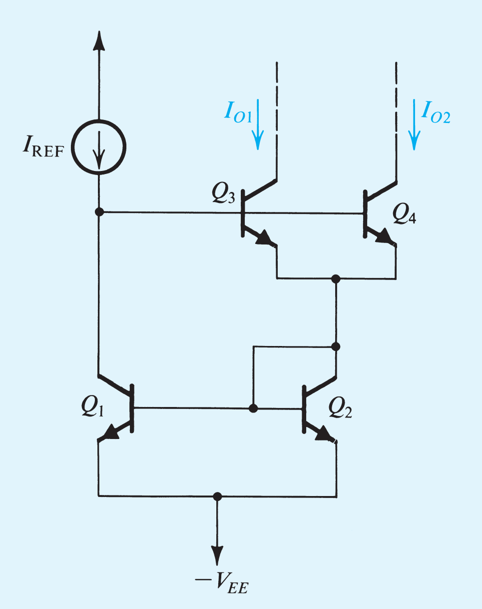


Figure P8.85

D 8.86 Use the *pnp* version of the Wilson current mirror to design a 0.1-mA current source. The current source is required to operate with the voltage at its output terminal as low as –2.5 V. If the power supplies available are ±2.5 V, what is the highest voltage possible at the output terminal?

***8.87** For the Wilson current mirror of Fig. 8.40, show that the incremental input resistance seen by I_{REF} is approximately $2V_T/I_{REF}$. (Neglect the Early effect in this derivation and assume a signal ground at the output.) Evaluate R_{in} for $I_{REF} =$ 0.2 mA.

***8.88** Consider the Wilson MOS mirror of Fig. 8.41(a) for the case of all transistors identical, with $W/L = 10$, $\mu_n C_{ox} =$ 400 μA/V^2, $V_{tn} = 0.5$ V, and $V_A = 18$ V. The mirror is fed with $I_{REF} = 180$ μA.

(a) Obtain an estimate of V_{OV} and V_{GS} at which the three transistors are operating, by neglecting the Early effect.
(b) Noting that Q_1 and Q_2 are operating at different V_{DS}, obtain an approximate value for the difference in their currents and hence determine I_O.
(c) To eliminate the systematic error between I_O and I_{REF} caused by the difference in V_{DS} between Q_1 and Q_2, a diode-connected transistor Q_4 can be added to the circuit

as shown in Fig. 8.41(c). What do you estimate I_O now to be?

(d) What is the minimum allowable voltage at the output node of the mirror?

(e) Convince yourself that Q_4 will have no effect on the output resistance of the mirror. Find R_o.

(f) What is the change in I_O (both absolute value and percentage) that results from $\Delta V_O = 1$ V?

8.89 Show that the incremental input resistance (seen by I_{REF}) for the Wilson MOS mirror of Fig. 8.41(a) is $2/g_m$. Assume that all three transistors are identical and neglect the Early effect. Also, assume a signal ground at the output. (*Hint*: Replace all transistors by their T model and remember that Q_1 is equivalent to a resistance $1/g_m$.)

D 8.90 (a) Utilizing a reference current of 200 μA, design a Widlar current source to provide an output current of 20 μA. Assume β to be high.

(b) If $\beta = 200$ and $V_A = 50$ V, find the value of the output resistance, and find the change in output current corresponding to a 5-V change in output voltage.

D 8.91 Design three Widlar current sources, each having a 100-μA reference current: one with a current transfer ratio of 0.8, one with a ratio of 0.10, and one with a ratio of 0.01, all assuming high β. For each, find the output resistance, and contrast it with r_o of the basic unity-ratio source that is providing the desired current and for which $R_E = 0$. Use $\beta = \infty$ and $V_A = 50$ V.

D 8.92 (a) For the circuit in Fig. P8.92, assume BJTs with high β and $v_{BE} = 0.7$ V at 1 mA. Find the value of R that will result in $I_O = 10$ μA.

(b) For the design in (a), find R_o assuming $\beta = 100$ and $V_A = 40$ V.

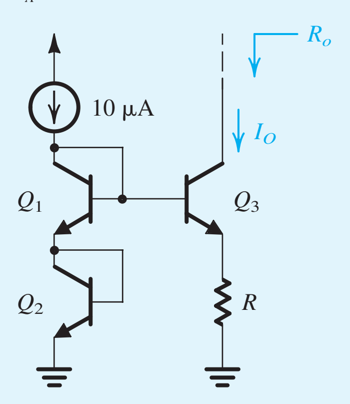


Figure P8.92

D 8.93 If the *pnp* transistor in the circuit of Fig. P8.93 is characterized by its exponential relationship with a scale current I_S, show that the dc current I is determined by $IR = V_T \ln(I/I_S)$. Assume Q_1 and Q_2 to be matched and Q_3, Q_4, and Q_5 to be matched. Find the value of R that yields a current $I = 200$ μA. For the BJT, $V_{EB} = 0.7$ V at $I_E = 1$ mA.

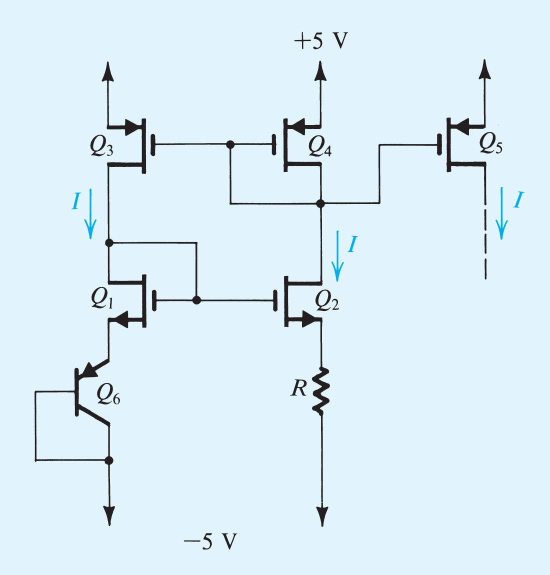


Figure P8.93

Section 8.7: Some Useful Transistor Pairings

8.94 Use the source-follower equivalent circuit in Fig. 8.45(b) to show that its output resistance is given by

$$R_o = r_{o3} \| r_{o1} \| \frac{1}{g_m + g_{mb}} \simeq \frac{1}{g_m + g_{mb}}$$

8.95 A source follower for which $k'_n = 200$ μA/V^2, $V'_A = 20$ V/μm, $\chi = 0.2$, $L = 0.5$ μm, $W = 20$ μm, and $V_t = 0.6$ V is required to provide a dc level shift (between input and output of 0.9 V.) What must the bias current be? Find g_m, g_{mb}, r_o, A_{vo}, and R_o. Assume that the bias current source has an output resistance equal to r_o. Also find the voltage gain when a load resistance of 2 kΩ is connected to the output.

8.96 The transistors in the circuit of Fig. P8.96 have $\beta = 100$ and $V_A = 50$ V.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

(a) Find R_{in} and the overall voltage gain.
(b) What is the effect of increasing the bias currents by a factor of 10 on R_{in}, G_v, and the power dissipation?

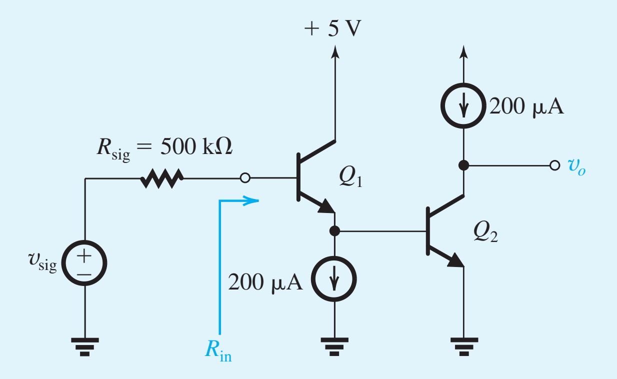


Figure P8.96

D *8.97 Consider the BiCMOS amplifier shown in Fig. P8.97. The BJT has $V_{BE} = 0.7$ V and $\beta = 200$. The MOSFET has $V_t = 1$ V and $k_n = 2$ mA/V^2. Neglect the Early effect in both devices.

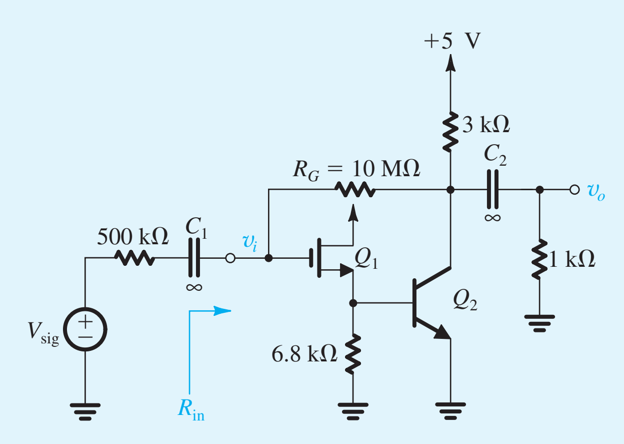


Figure P8.97

(a) Consider the dc bias circuit. Neglect the base current in Q_2 in determining the current in Q_1. Find the dc bias currents in Q_1 and Q_2 and show that they are approximately 100 μA and 1 mA, respectively.
(b) Evaluate the small-signal parameters of Q_1 and Q_2 at their bias points.
(c) Determine the voltage gain $A_v = v_o/v_i$. For this purpose you can neglect R_G.
(d) Noting that R_G is connected between the input node where the voltage is v_i and the output node where the voltage is $A_v v_i$, find R_{in} and hence the overall voltage gain v_o/v_{sig}.
(e) To considerably reduce the effect of R_G on R_{in} and hence on G_v, consider the effect of adding another 10-MΩ resistor in series with the existing one and placing a large bypass capacitor between their joint node and ground. What will R_{in} and G_v become?

8.98 The BJTs in the Darlington follower of Fig. P8.98 have $\beta = 100$. If the follower is fed with a source having a 100-kΩ resistance and is loaded with 1 kΩ, find the input resistance and the output resistance (excluding the load). Also find the overall voltage gain, both open-circuited and with load. Neglect the Early effect.

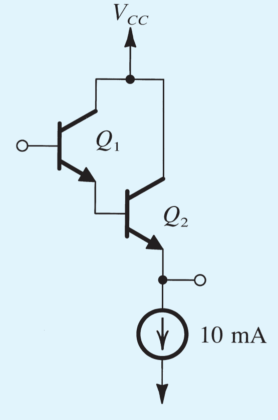


Figure P8.98

8.99 For the amplifier in Fig. 8.48(a), let $I = 0.5$ mA and $\beta = 100$, and neglect r_o. Assume that a load resistance of 10 kΩ is connected to the output terminal. If the amplifier is fed with a signal v_{sig} having a source resistance $R_{sig} = 10$ kΩ, find G_v.

8.100 Consider the CD–CG amplifier of Fig. 8.48(c) for the case $g_m = 5$ mA/V, $R_{sig} = 500$ kΩ, and $R_L = 10$ kΩ. Neglecting r_o, find G_v.

****8.101** In each of the six circuits in Fig. P8.101, let $\beta = 100$, and neglect r_o. Calculate the overall voltage gain.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

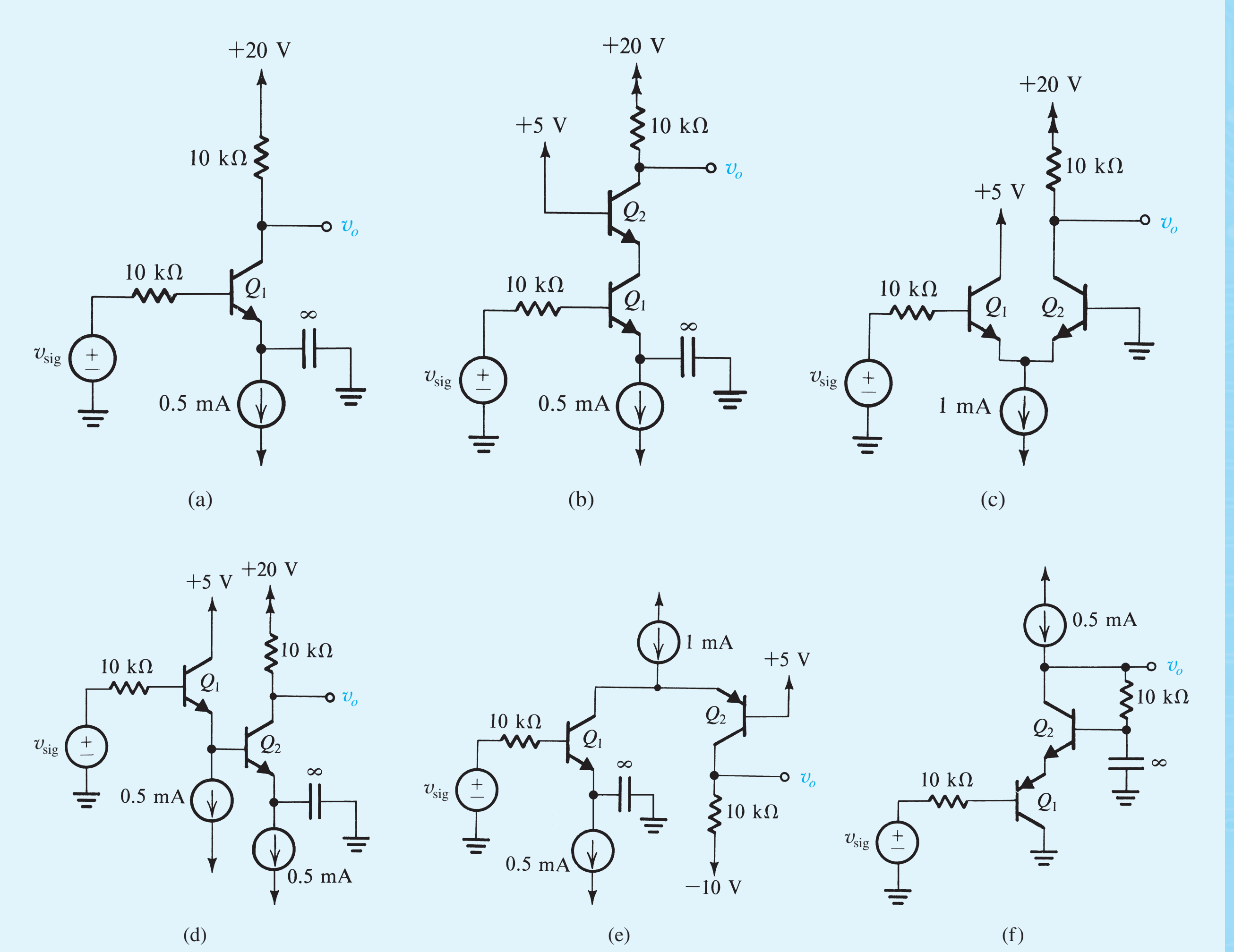


Figure P8.101

CHAPTER 9

Differential and Multistage Amplifiers

IN THIS CHAPTER YOU WILL LEARN

1. The essence of the operation of the MOS and the bipolar differential amplifiers: how they reject common-mode noise or interference and amplify differential signals.
2. The analysis and design of MOS and BJT differential amplifiers.
3. Differential-amplifier circuits of varying complexity; utilizing passive resistive loads, current-source loads, and cascodes—the building blocks we studied in Chapter 8.
4. An ingenious and highly popular differential-amplifier circuit that utilizes a current-mirror load.
5. The structure, analysis, and design of amplifiers composed of two or more stages in cascade. Two practical examples are studied in detail: a two-stage CMOS op amp and a four-stage bipolar op amp.

Introduction

The differential-pair or differential-amplifier configuration is the most widely used building block in analog integrated-circuit design. For instance, the input stage of every op amp is a differential amplifier. Also, the BJT differential amplifier is the basis of a very-high-speed logic-circuit family, studied in Chapter 15, called emitter-coupled logic (ECL).

Initially invented in the 1940s for use with vacuum tubes, the basic differential-amplifier configuration was subsequently implemented with discrete bipolar transistors. However, it was the advent of integrated circuits that has made the differential pair extremely popular in both bipolar and MOS technologies. There are two reasons why differential amplifiers are so well suited for IC fabrication: First, as we shall shortly see, the performance of the differential pair depends critically on the matching between the two sides of the circuit. Integrated-circuit fabrication is capable of providing matched devices whose parameters track over wide ranges of changes in environmental conditions. Second, by their very nature, differential amplifiers utilize more components (approaching twice as many) than single-ended circuits. Here again, the reader will recall from the discussion in Section 8.1 that a significant advantage of integrated-circuit technology is the availability of large numbers of transistors at relatively low cost.

We assume that the reader is familiar with the basic concept of a differential amplifier as presented in Section 2.1. Nevertheless it is worthwhile to answer the question: Why differential? Basically, there are two reasons for using differential in preference to single-ended amplifiers. First, differential circuits are much less sensitive to noise and interference than

single-ended circuits. To appreciate this point, consider two wires carrying a small differential signal as the voltage difference between the two wires. Now, assume that there is an interference signal that is coupled to the two wires, either capacitively or inductively. As the two wires are physically close together, the interference voltages on the two wires (i.e., between each of the two wires and ground) will be equal. Since, in a differential system, only the difference signal between the two wires is sensed, it will contain no interference component!

The second reason for preferring differential amplifiers is that the differential configuration enables us to bias the amplifier and to couple amplifier stages together without the need for bypass and coupling capacitors such as those utilized in the design of discrete-circuit amplifiers (Section 7.5). This is another reason why differential circuits are ideally suited for IC fabrication where large capacitors are impossible to fabricate economically.

The major topic of this chapter is the differential amplifier in both its MOS and bipolar implementations. As will be seen, the design and analysis of differential amplifiers makes extensive use of the material on single-stage amplifiers presented in Chapters 7 and 8. We will follow the study of differential amplifiers with examples of practical multistage amplifiers, again in both MOS and bipolar technologies.

9.1 The MOS Differential Pair

Figure 9.1 shows the basic MOS differential-pair configuration. It consists of two matched transistors, Q_1 and Q_2, whose sources are joined together and biased by a constant-current source I. The latter is usually implemented by a MOSFET circuit of the type studied in Sections 8.2 and 8.5. For the time being, we assume that the current source is ideal and that it has infinite output resistance. Although each drain is shown connected to the positive supply through a resistance R_D, in most cases active (current-source) loads are employed, as will be seen shortly. For the time being, however, we will explain the essence of the differential-pair operation utilizing simple resistive loads. Whatever type of load is used, it is essential that the MOSFETs not enter the triode region of operation.

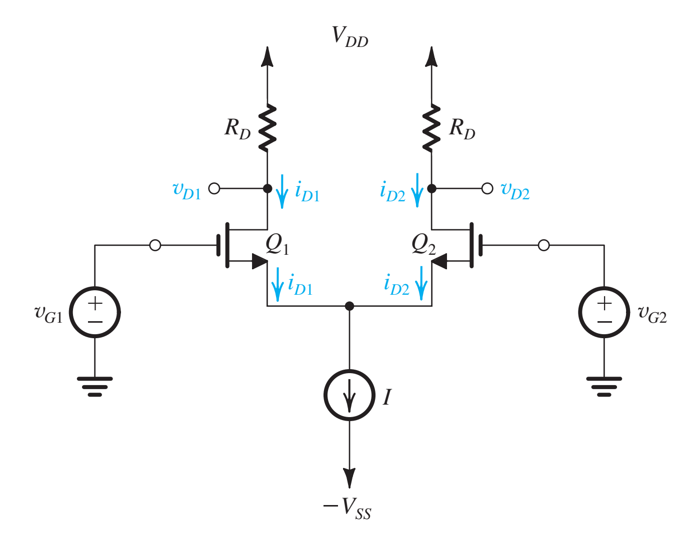


Figure 9.1 The basic MOS differential-pair configuration.

9.1.1 Operation with a Common-Mode Input Voltage

To see how the differential pair works, consider first the case when equal voltages, V_{CM}, are applied to the two gate terminals. That is, as shown in Fig. 9.2, $v_{G1} = v_{G2} = V_{CM}$. Since V_{CM} is common to the two input terminals, it is called **common-mode voltage**. Since Q_1 and Q_2 are matched, the current I will divide equally between the two transistors. Thus, $i_{D1} = i_{D2} = I/2$, and the voltage at the sources, V_S, will be

$$V_S = V_{CM} - V_{GS} \tag{9.1}$$

where V_{GS} is the gate-to-source voltage corresponding to a drain current of $I/2$. Neglecting channel-length modulation, V_{GS} and $I/2$ are related by

$$\frac{I}{2} = \frac{1}{2}k_n'\frac{W}{L}(V_{GS} - V_t)^2 \tag{9.2}$$

or in terms of the overdrive voltage V_{OV},

$$V_{OV} = V_{GS} - V_t \tag{9.3}$$

$$\frac{I}{2} = \frac{1}{2}k_n'\frac{W}{L}V_{OV}^2 \tag{9.4}$$

$$V_{OV} = \sqrt{I/k_n'(W/L)} \tag{9.5}$$

The voltage at each drain will be

$$v_{D1} = v_{D2} = V_{DD} - \frac{I}{2}R_D \tag{9.6}$$

Thus, the difference in voltage between the two drains will be zero.

Now, let us vary the value of the common-mode voltage V_{CM}. We see that, as long as Q_1 and Q_2 remain in the saturation region, the current I will divide equally between Q_1 and Q_2 and the voltages at the drains will not change. Thus the differential pair does *not* respond to (i.e., it *rejects*) common-mode input signals.

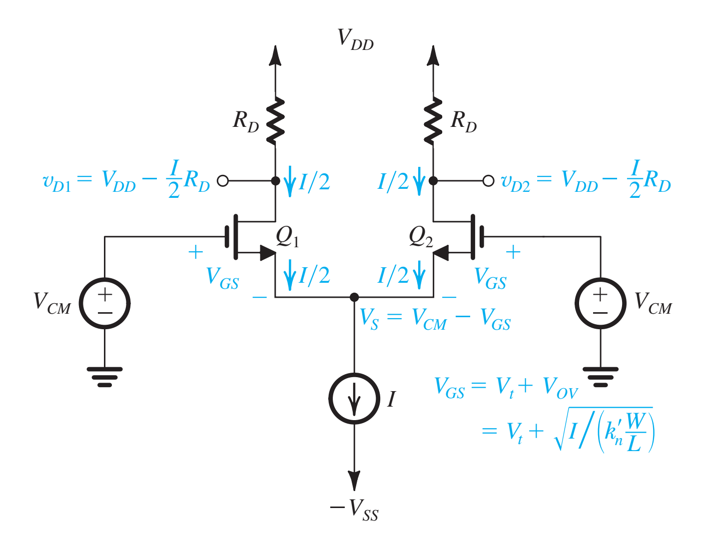


Figure 9.2 The MOS differential pair with a common-mode input voltage V_{CM}.

An important specification of a differential amplifier is its **input common-mode range**. This is the range of V_{CM} over which the differential pair operates properly. The highest value of V_{CM} is limited by the requirement that Q_1 and Q_2 remain in saturation, thus

$$V_{CM\max} = V_t + V_{DD} - \frac{I}{2}R_D \tag{9.7}$$

The lowest value of V_{CM} is determined by the need to allow for a sufficient voltage across the current source I for it to operate properly. If a voltage V_{CS} is needed across the current source, then

$$V_{CM\min} = -V_{SS} + V_{CS} + V_t + V_{OV} \tag{9.8}$$

Example 9.1

For the MOS differential pair with a common-mode voltage V_{CM} applied, as shown in Fig. 9.2, let $V_{DD} = V_{SS} = 1.5$ V, $k_n'(W/L) = 4$ mA/V^2, $V_t = 0.5$ V, $I = 0.4$ mA, and $R_D = 2.5$ kΩ, and neglect channel-length modulation. Assume that the current source I requires a minimum voltage of 0.4 V to operate properly.

(a) Find V_{OV} and V_{GS} for each transistor.
(b) For $V_{CM} = 0$, find V_S, I_{D1}, I_{D2}, V_{D1}, and V_{D2}.
(c) Repeat (b) for $V_{CM} = +1$ V.
(d) Repeat (b) for $V_{CM} = -0.2$ V.
(e) What is the highest permitted value of V_{CM}?
(f) What is the lowest value allowed for V_{CM}?

Solution

(a) With $v_{G1} = v_{G2} = V_{CM}$, we see that $V_{GS1} = V_{GS2}$. Now, since the transistors are matched, I will divide equally between the two transistors,

$$I_{D1} = I_{D2} = \frac{I}{2}$$

Thus,

$$\frac{I}{2} = \frac{1}{2}k_n'(W/L)V_{OV}^2$$

$$\frac{0.4}{2} = \frac{1}{2} \times 4V_{OV}^2$$

which results in

$$V_{OV} = 0.316 \text{ V}$$

and thus,

$$V_{GS} = V_t + V_{OV} = 0.5 + 0.316 \simeq 0.82 \text{ V}$$

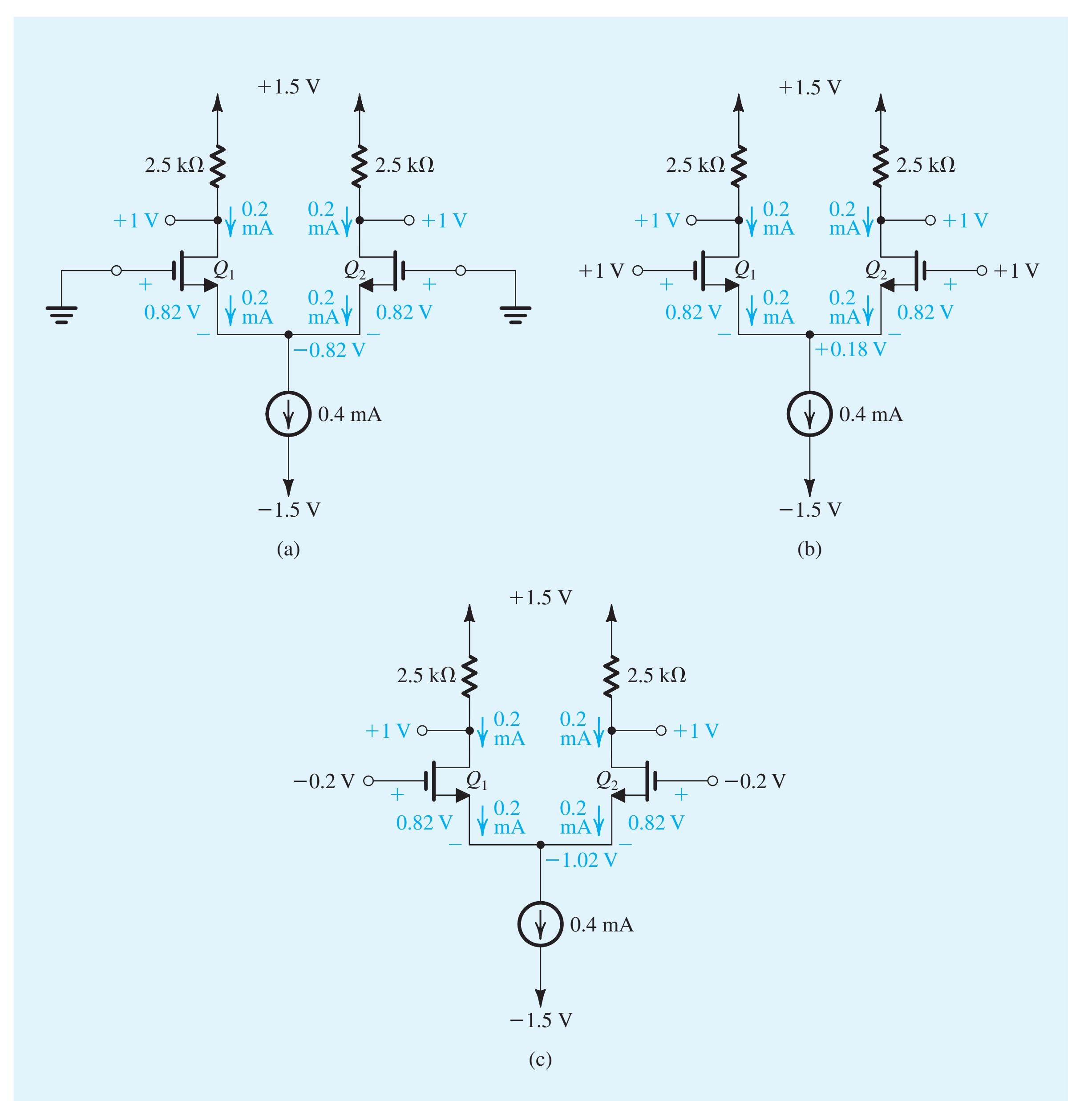


Figure 9.3 Circuits for Example 9.1. Effects of varying V_{CM} on the operation of the differential pair.

(b) The analysis for the case $V_{CM} = 0$ is shown in Fig. 9.3(a) from which we see that

$$V_S = V_G - V_{GS} = 0 - 0.82 = -0.82 \text{ V}$$

$$I_{D1} = I_{D2} = \frac{I}{2} = 0.2 \text{ mA}$$

Example 9.1 *continued*

$$V_{D1} = V_{D2} = V_{DD} - \frac{I}{2}R_D$$
$$= 1.5 - 0.2 \times 2.5 = 1 \text{ V}$$

(c) The analysis for the case $V_{CM} = +1$ V is shown in Fig. 9.3(b) from which we see that

$$V_S = V_G - V_{GS} = 1 - 0.82 = +0.18 \text{ V}$$
$$I_{D1} = I_{D2} = \frac{I}{2} = 0.2 \text{ mA}$$
$$V_{D1} = V_{D2} = V_{DD} - \frac{I}{2}R_D = 1.5 - 0.2 \times 2.5 = +1 \text{ V}$$

Observe that the transistors remain in the saturation region as assumed. Also observe that I_{D1}, I_{D2}, V_{D1}, and V_{D2} remain unchanged even though the common-mode voltage V_{CM} changed by 1 V.

(d) The analysis for the case $V_{CM} = -0.2$ V is shown in Fig. 9.3(c), from which we see that

$$V_S = V_G - V_{GS} = -0.2 - 0.82 = -1.02 \text{ V}$$

It follows that the current source I now has a voltage across it of

$$V_{CS} = -V_S - (-V_{SS}) = -1.02 + 1.5 = 0.48 \text{ V}$$

which is greater than the minimum required value of 0.4 V. Thus, the current source is still operating properly and delivering a constant current $I = 0.4$ mA and hence

$$I_{D1} = I_{D2} = \frac{I}{2} = 0.2 \text{ mA}$$
$$V_{D1} = V_{D2} = V_{DD} - \frac{I}{2}R_D = +1 \text{ V}$$

So, here again the differential circuit is not responsive to the change in the common-mode voltage V_{CM}.

(e) The highest permitted value of V_{CM} is that which causes Q_1 and Q_2 to leave saturation and enter the triode region. Thus,

$$V_{CM\max} = V_t + V_D$$
$$= 0.5 + 1 = +1.5 \text{ V}$$

(f) The lowest value allowed for V_{CM} is that which reduces the voltage across the current source I to the minimum required of $V_{CS} = 0.4$ V. Thus,

$$V_{CM\min} = -V_{SS} + V_{CS} + V_{GS}$$
$$= -1.5 + 0.4 + 0.82 = -0.28 \text{ V}$$

Thus, the input common-mode range is

$$-0.28 \text{ V} \leq V_{CM} \leq +1.5 \text{ V}$$

EXERCISE

9.1 For the amplifier in Example 9.1, find the input common-mode range for the case in which the two drain resistances R_D are increased by a factor of 2.
Ans. −0.28 V to 1.0 V

9.1.2 Operation with a Differential Input Voltage

Next we apply a difference or differential input voltage by grounding the gate of Q_2 (i.e., setting $v_{G2} = 0$) and applying a signal v_{id} to the gate of Q_1, as shown in Fig. 9.4. We can see that since $v_{id} = v_{GS1} - v_{GS2}$, if v_{id} is positive, v_{GS1} will be greater than v_{GS2} and hence i_{D1} will be greater than i_{D2} and the difference output voltage $(v_{D2} - v_{D1})$ will be positive. On the other hand, when v_{id} is negative, v_{GS1} will be lower than v_{GS2}, i_{D1} will be smaller than i_{D2}, and correspondingly v_{D1} will be higher than v_{D2}; in other words, the difference or differential output voltage $(v_{D2} - v_{D1})$ will be negative.

From the above, we see that the differential pair responds to **difference-mode** or **differential input signals** by providing a corresponding differential output signal between the two drains. At this point, it is useful to inquire about the value of v_{id} that causes the entire bias current I to flow in one of the two transistors. In the positive direction, this happens when v_{GS1} reaches the value that corresponds to $i_{D1} = I$, and v_{GS2} is reduced to a value equal to the threshold voltage V_t, at which point $v_S = -V_t$. The value of v_{GS1} can be found from

$$I = \frac{1}{2}\left(k_n' \frac{W}{L}\right)(v_{GS1} - V_t)^2$$

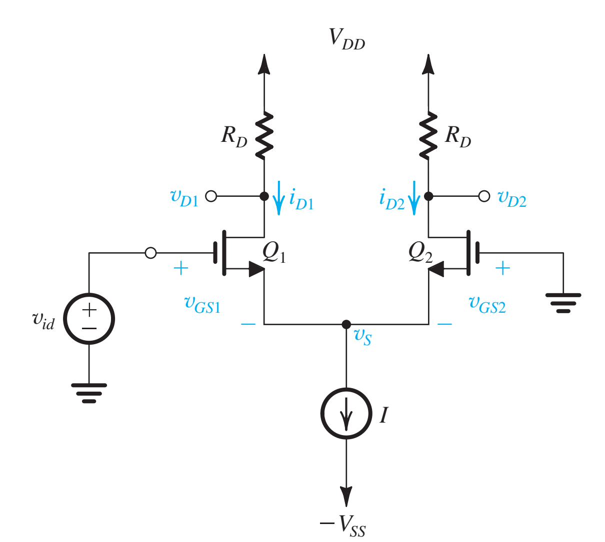


Figure 9.4 The MOS differential pair with a differential input signal v_{id} applied. With v_{id} positive: $v_{GS1} > v_{GS2}$, $i_{D1} > i_{D2}$, and $v_{D1} < v_{D2}$; thus $(v_{D2} - v_{D1})$ will be positive. With v_{id} negative: $v_{GS1} < v_{GS2}$, $i_{D1} < i_{D2}$, and $v_{D1} > v_{D2}$; thus $(v_{D2} - v_{D1})$ will be negative.

as

$$\begin{aligned} v_{GS1} &= V_t + \sqrt{2I/k_n'(W/L)} \\ &= V_t + \sqrt{2}V_{OV} \end{aligned} \tag{9.9}$$

where V_{OV} is the overdrive voltage corresponding to a drain current of $I/2$ (Eq. 9.5). Thus, the value of v_{id} at which the entire bias current I is steered into Q_1 is

$$\begin{aligned} v_{id\max} &= v_{GS1} + v_S \\ &= V_t + \sqrt{2}V_{OV} - V_t \\ &= \sqrt{2}V_{OV} \end{aligned} \tag{9.10}$$

If v_{id} is increased beyond $\sqrt{2}V_{OV}$, i_{D1} remains equal to I, v_{GS1} remains equal to $(V_t + \sqrt{2}V_{OV})$, and v_S rises correspondingly, thus keeping Q_2 off. In a similar manner we can show that in the negative direction, as v_{id} reaches $-\sqrt{2}V_{OV}$, Q_1 turns off and Q_2 conducts the entire bias current I. Thus the current I can be steered from one transistor to the other by varying v_{id} in the range



$$-\sqrt{2}V_{OV} \leq v_{id} \leq \sqrt{2}V_{OV}$$

which defines the range of differential-mode operation. Finally, observe that we have assumed that Q_1 and Q_2 remain in saturation even when one of them is conducting the entire current I.

EXERCISE

9.2 For the MOS differential pair specified in Example 9.1 find (a) the value of v_{id} that causes Q_1 to conduct the entire current I, and the corresponding values of v_{D1} and v_{D2}; (b) the value of v_{id} that causes Q_2 to conduct the entire current I, and the corresponding values of v_{D1} and v_{D2}; (c) the corresponding range of the differential output voltage $(v_{D2} - v_{D1})$.
Ans. (a) +0.45 V, 0.5 V, 1.5 V; (b) –0.45 V, 1.5 V, 0.5 V; (c) +1 V to –1 V

To use the differential pair as a linear amplifier, we keep the differential input signal v_{id} small. As a result, the current in one of the transistors (Q_1 when v_{id} is positive) will increase by an increment ΔI proportional to v_{id}, to $(I/2 + \Delta I)$. Simultaneously, the current in the other transistor will decrease by the same amount to become $(I/2 - \Delta I)$. A voltage signal $-\Delta IR_D$ develops at one of the drains and an opposite-polarity signal, ΔIR_D, develops at the other drain. Thus the output voltage taken between the two drains will be $2\Delta I\, R_D$, which is proportional to the differential input signal v_{id}. The small-signal operation of the differential pair will be studied in detail in Section 9.1.4.

9.1.3 Large-Signal Operation

We shall now derive expressions for the drain currents i_{D1} and i_{D2} in terms of the input differential signal $v_{id} \equiv v_{G1} - v_{G2}$. The derivation assumes that the differential pair is perfectly

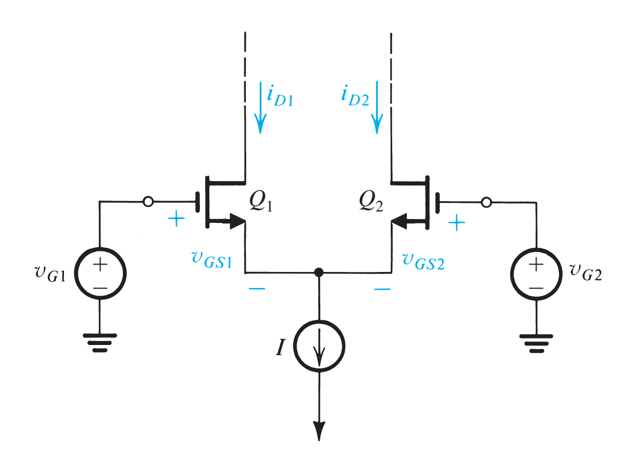


Figure 9.5 The MOSFET differential pair for the purpose of deriving the transfer characteristics, i_{D1} and i_{D2} versus $v_{id} = v_{G1} - v_{G2}$.

matched and neglects channel-length modulation ($\lambda = 0$ or $V_A = \infty$). Thus these expressions do not depend on the details of the circuit to which the drains are connected, and we do not show these connections in Fig. 9.5; we simply assume that the circuit maintains Q_1 and Q_2 in the saturation region of operation at all times.

To begin with, we express the drain currents of Q_1 and Q_2 as

$$i_{D1} = \frac{1}{2}k_n'\frac{W}{L}(v_{GS1} - V_t)^2 \tag{9.11}$$

$$i_{D2} = \frac{1}{2}k_n'\frac{W}{L}(v_{GS2} - V_t)^2 \tag{9.12}$$

Taking the square roots of both sides of each of Eqs. (9.11) and (9.12), we obtain

$$\sqrt{i_{D1}} = \sqrt{\frac{1}{2}k_n'\frac{W}{L}}(v_{GS1} - V_t) \tag{9.13}$$

$$\sqrt{i_{D2}} = \sqrt{\frac{1}{2}k_n'\frac{W}{L}}(v_{GS2} - V_t) \tag{9.14}$$

Subtracting Eq. (9.14) from Eq. (9.13) and substituting

$$v_{GS1} - v_{GS2} = v_{G1} - v_{G2} = v_{id} \tag{9.15}$$

results in

$$\sqrt{i_{D1}} - \sqrt{i_{D2}} = \sqrt{\frac{1}{2}k_n'\frac{W}{L}}\,v_{id} \tag{9.16}$$

The constant-current bias imposes the constraint

$$i_{D1} + i_{D2} = I \tag{9.17}$$

Equations (9.16) and (9.17) are two equations in the two unknowns i_{D1} and i_{D2} and can be solved as follows: Squaring both sides of Eq. (9.16) and substituting for $i_{D1} + i_{D2} = I$ gives

$$2\sqrt{i_{D1}i_{D2}} = I - \frac{1}{2}k_n'\frac{W}{L}v_{id}^2$$

Substituting for i_{D2} from Eq. (9.17) as $i_{D2} = I - i_{D1}$ and squaring both sides of the resulting equation provides a quadratic equation in i_{D1} that can be solved to yield

$$i_{D1} = \frac{I}{2} \pm \sqrt{k_n' \frac{W}{L} I} \left(\frac{v_{id}}{2}\right) \sqrt{1 - \frac{(v_{id}/2)^2}{I/k_n' \frac{W}{L}}}$$

Now, since the increment in i_{D1} above the bias value of $I/2$ must have the same polarity as v_{id}, only the root with the "+" sign in the second term is physically meaningful; thus,

$$i_{D1} = \frac{I}{2} + \sqrt{k_n' \frac{W}{L} I} \left(\frac{v_{id}}{2}\right) \sqrt{1 - \frac{(v_{id}/2)^2}{I/k_n' \frac{W}{L}}} \tag{9.18}$$

The corresponding value of i_{D2} is found from $i_{D2} = I - i_{D1}$ as

$$i_{D2} = \frac{I}{2} - \sqrt{k_n' \frac{W}{L} I} \left(\frac{v_{id}}{2}\right) \sqrt{1 - \frac{(v_{id}/2)^2}{I/k_n' \frac{W}{L}}} \tag{9.19}$$

At the bias (quiescent) point, $v_{id} = 0$, leading to

$$i_{D1} = i_{D2} = \frac{I}{2} \tag{9.20}$$

Correspondingly,

$$v_{GS1} = v_{GS2} = V_{GS} \tag{9.21}$$

where

$$\frac{I}{2} = \frac{1}{2} k_n' \frac{W}{L} (V_{GS} - V_t)^2 = \frac{1}{2} k_n' \frac{W}{L} V_{OV}^2 \tag{9.22}$$

This relationship enables us to replace $k_n'(W/L)$ in Eqs. (9.18) and (9.19) with I/V_{OV}^2 to express i_{D1} and i_{D2} in the alternative form

$$i_{D1} = \frac{I}{2} + \left(\frac{I}{V_{OV}}\right)\left(\frac{v_{id}}{2}\right)\sqrt{1 - \left(\frac{v_{id}/2}{V_{OV}}\right)^2} \tag{9.23}$$

$$i_{D2} = \frac{I}{2} - \left(\frac{I}{V_{OV}}\right)\left(\frac{v_{id}}{2}\right)\sqrt{1 - \left(\frac{v_{id}/2}{V_{OV}}\right)^2} \tag{9.24}$$

These two equations describe the effect of applying a differential input signal v_{id} on the currents i_{D1} and i_{D2}. They can be used to obtain the normalized plots, i_{D1}/I and i_{D2}/I versus v_{id}/V_{OV}, shown in Fig. 9.6. Note that at $v_{id} = 0$, the two currents are equal to $I/2$. Making v_{id} positive causes i_{D1} to increase and i_{D2} to decrease by equal amounts, to keep the sum constant, $i_{D1} + i_{D2} = I$. The current is steered entirely into Q_1 when v_{id} reaches the value $\sqrt{2}V_{OV}$, as we found out earlier. For v_{id} negative, identical statements can be made by interchanging i_{D1} and i_{D2}. In this case, $v_{id} = -\sqrt{2}V_{OV}$ steers the current entirely into Q_2. Finally, note that the plots in Fig. 9.6 are universal, as they apply to any MOS differential pair.

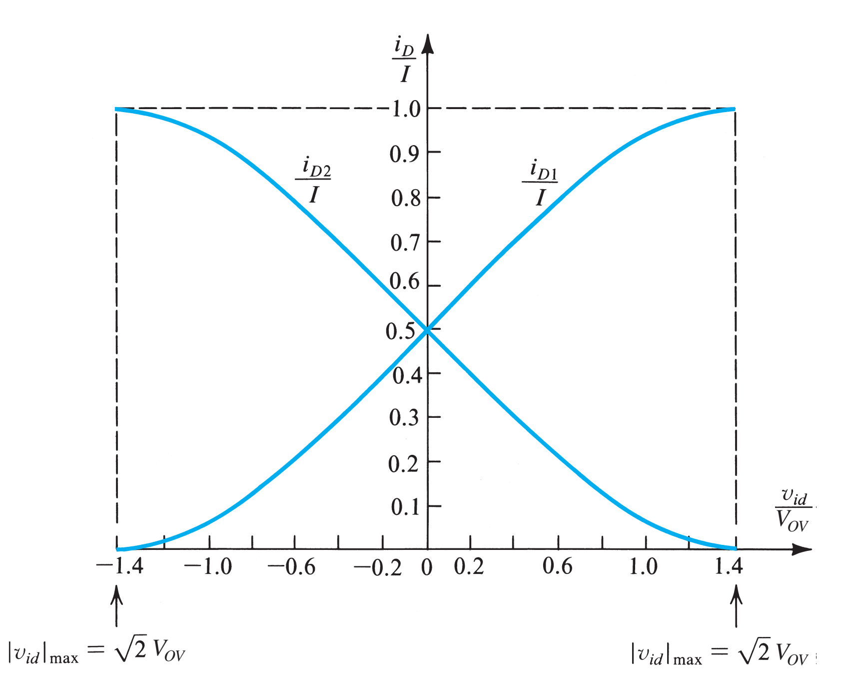


Figure 9.6 Normalized plots of the currents in a MOSFET differential pair. Note that V_{OV} is the overdrive voltage at which Q_1 and Q_2 operate when conducting drain currents equal to $I/2$, the equilibrium situation. Note that these graphs are universal and apply to any MOS differential pair.

The transfer characteristics of Eqs. (9.23) and (9.24) and Fig. 9.6 are obviously nonlinear. This is due to the term involving v_{id}^2. Since we are interested in obtaining linear amplification from the differential pair, we will strive to make this term as small as possible. For a given value of V_{OV}, the only thing we can do is keep $(v_{id}/2)$ much smaller than V_{OV}, which is the condition for the small-signal approximation. It results in

$$i_{D1} \simeq \frac{I}{2} + \left(\frac{I}{V_{OV}}\right)\left(\frac{v_{id}}{2}\right) \tag{9.25}$$

and

$$i_{D2} \simeq \frac{I}{2} - \left(\frac{I}{V_{OV}}\right)\left(\frac{v_{id}}{2}\right) \tag{9.26}$$

which, as expected, indicate that i_{D1} increases by an increment i_d, and i_{D2} decreases by the same amount, i_d, where i_d is proportional to the differential input signal v_{id},

$$i_d = \left(\frac{I}{V_{OV}}\right)\left(\frac{v_{id}}{2}\right) \tag{9.27}$$

Recalling from our study of the MOSFET amplifier in Chapter 7 (also refer to Table G.3 in Appendix G), that a MOSFET biased at a current I_D has a transconductance $g_m = 2I_D/V_{OV}$, we recognize the factor (I/V_{OV}) in Eq. (9.27) as g_m of each of Q_1 and Q_2, which are biased at $I_D = I/2$. Now, why $v_{id}/2$? Simply because v_{id} divides equally between the two devices with $v_{gs1} = v_{id}/2$ and $v_{gs2} = -v_{id}/2$, which causes Q_1 to have a current increment i_d and Q_2 to have a current decrement i_d. We shall analyze the small-signal operation of the MOS differential pair

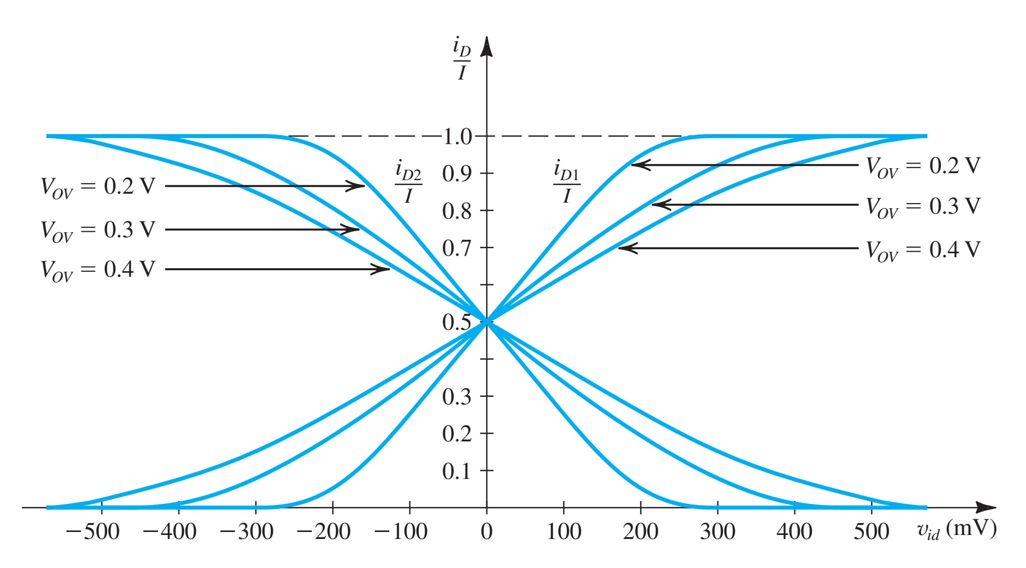


Figure 9.7 The linear range of operation of the MOS differential pair can be extended by operating the transistor at a higher value of V_{OV}.

shortly. At this time, however, we wish to return to Eqs. (9.23) and (9.24) and note that for a given v_{id}, linearity can be increased by increasing the overdrive voltage V_{OV} at which each of Q_1 and Q_2 is operating. This can be done by using smaller W/L ratios. The price paid for the increased linearity is a reduction in g_m and hence a reduction in gain. In this regard, we observe that the normalized plot of Fig. 9.6, though compact, masks this design degree of freedom. Figure 9.7 shows plots of the transfer characteristics $i_{D1,2}/I$ versus v_{id} for various values of V_{OV}. These graphs clearly illustrate the linearity–transconductance trade-off obtained by changing the value of V_{OV}: The linear range of operation can be extended by operating the MOSFETs at a higher V_{OV} (by using smaller W/L ratios) at the expense of reducing g_m and hence the gain. This trade-off is based on the assumption that the bias current I is being kept constant. The bias current can, of course, be increased to obtain a higher g_m. The expense for doing this, however, is increased power dissipation, a serious limitation in IC design.

EXERCISE

9.3 A MOS differential pair is operated at a bias current I of 0.4 mA. If $\mu_n C_{ox} = 0.2\ \text{mA/V}^2$, find the required values of W/L and the resulting g_m if the MOSFETs are operated at $V_{OV} = 0.2$, 0.3, and 0.4 V. For each value, give the maximum $|v_{id}|$ for which the term involving v_{id}^2 in Eqs. (9.23) and (9.24), namely, $((v_{id}/2)/V_{OV})^2$, is limited to 0.1.

Ans.

V_{OV} (V)	0.2	0.3	0.4
W/L	50	22.2	12.5
g_m (mA/V)	2	1.33	1
$\lvert v_{id}\rvert_{max}$ (mV)	126	190	253

9.1.4 Small-Signal Operation

In this section we build on the understanding gained of the basic operation of the differential pair and consider in some detail its operation as a linear amplifier.

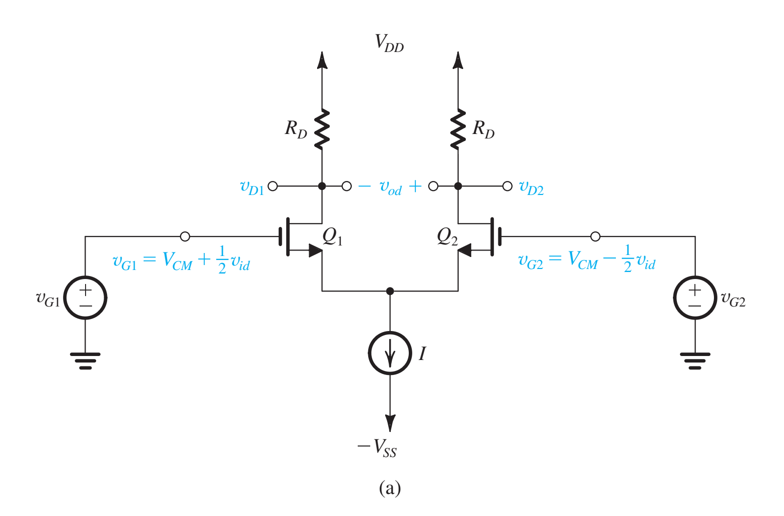
Differential Gain Figure 9.8(a) shows the MOS differential amplifier with input voltages

$$v_{G1} = V_{CM} + \frac{1}{2}v_{id} \tag{9.28}$$

and

$$v_{G2} = V_{CM} - \frac{1}{2}v_{id} \tag{9.29}$$

Here, V_{CM} denotes a common-mode dc voltage within the input common-mode range of the differential amplifier. It is needed in order to set the dc voltage of the MOSFET gates. Typically



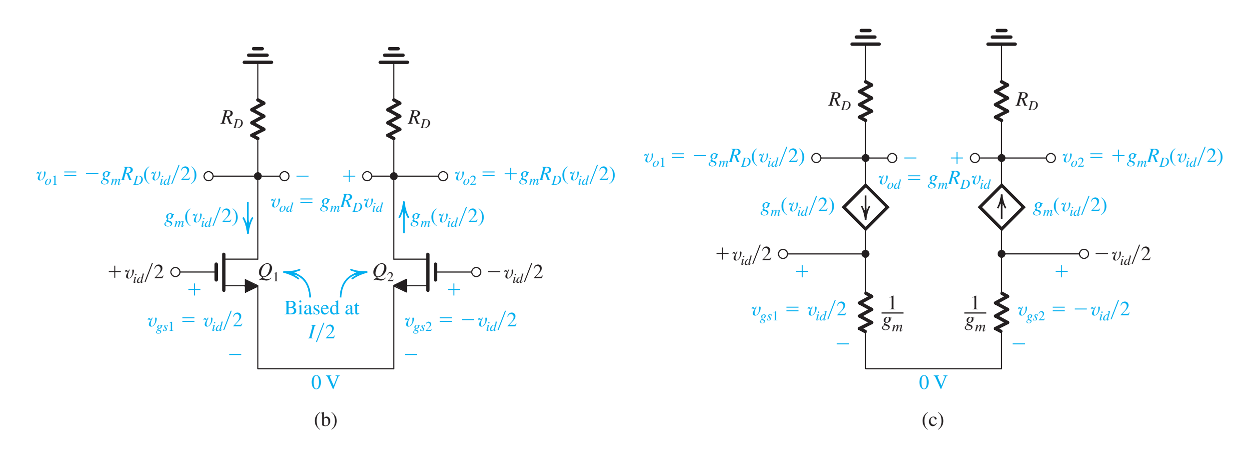


Figure 9.8 Small-signal analysis of the MOS differential amplifier. (**a**) The circuit with a common-mode voltage applied to set the dc bias voltage at the gates and with v_{id} applied in a complementary (or balanced) manner. (**b**) The circuit prepared for small-signal analysis. (**c**) The circuit in (**b**), with the MOSFETs replaced with T models.

V_{CM} is at the middle value of the power supply. Thus, for our case, where two complementary supplies are utilized, V_{CM} is typically 0 V.

The differential input signal v_{id} is applied in a **complementary** (or **balanced**) manner; that is, v_{G1} is increased by $v_{id}/2$ and v_{G2} is decreased by $v_{id}/2$. This would be the case, for instance, if the differential amplifier were fed from the output of another differential-amplifier stage. Sometimes, however, the differential input is applied in a single-ended fashion, as we saw earlier in Fig. 9.4. The difference in the performance resulting is too subtle a point for our current needs.

As indicated in Fig. 9.8(a) the amplifier output can be taken either between one of the drains and ground or between the two drains. In the first case, the resulting **single-ended outputs** v_{o1} and v_{o2} will be riding on top of the dc voltages at the drains, $\left(V_{DD} - \frac{I}{2}R_D\right)$. This is not the case when the output is taken between the two drains; the resulting **differential** output v_{od} (having a 0-V dc component) will be entirely a signal component. We will see shortly that there are other significant advantages to taking the output voltage differentially.

Our objective now is to analyze the small-signal operation of the differential amplifier of Fig. 9.8(a) to determine its voltage gain in response to the differential input signal v_{id}. Toward that end we show in Fig. 9.8(b) the circuit with the power supplies grounded, the bias current source I removed, and V_{CM} eliminated; that is, only signal quantities are indicated. For the time being we will neglect the effect of the MOSFET r_o. Finally note that each of Q_1 and Q_2 is biased at a dc current of $I/2$ and is operating at an overdrive voltage V_{OV}.

From the symmetry of the circuit and because of the balanced manner in which v_{id} is applied, we observe that the signal voltage at the joint source connection must be zero, acting as a sort of **virtual ground**. Thus Q_1 has a gate-to-source voltage signal $v_{gs1} = v_{id}/2$ and Q_2 has $v_{gs2} = -v_{id}/2$. Assuming $v_{id}/2 \ll V_{OV}$, the condition for the small-signal approximation, the changes resulting in the drain currents of Q_1 and Q_2 will be proportional to v_{gs1} and v_{gs2}, respectively. Thus Q_1 will have a drain current increment $g_m(v_{id}/2)$ and Q_2 will have a drain current decrement $g_m(v_{id}/2)$, where g_m denotes the equal transconductances of the two devices,

$$g_m = \frac{2I_D}{V_{OV}} = \frac{2(I/2)}{V_{OV}} = \frac{I}{V_{OV}} \tag{9.30}$$

These results correspond to those obtained earlier using the large-signal transfer characteristics and imposing the small-signal condition, Eqs. (9.25) to (9.27). To further illustrate the small-signal operation of the differential amplifier, we show in Fig. 9.8(c) its equivalent circuit obtained by replacing each of the MOSFETs with the corresponding T model. The reader is urged to study the correspondence between the elements and qualities in Fig. 9.8(b) and 9.8(c).

It is useful at this point to observe again that a signal ground is established at the source terminals of the transistors *without resorting to the use of a large bypass capacitor*, clearly a major advantage of the differential-pair configuration.

The essence of differential-pair operation is that it provides complementary current signals in the drains; what we do with the resulting pair of complementary current signals is, in a sense, a separate issue. Here, of course, we are simply passing the two current signals through a pair of matched resistors, R_D, and thus obtaining the drain voltage signals

$$v_{o1} = -g_m \frac{v_{id}}{2} R_D \tag{9.31}$$

and

$$v_{o2} = +g_m \frac{v_{id}}{2} R_D \tag{9.32}$$

If the output is taken in a single-ended fashion, the resulting gain becomes

$$\frac{v_{o1}}{v_{id}} = -\frac{1}{2} g_m R_D \tag{9.33}$$

or

$$\frac{v_{o2}}{v_{id}} = \frac{1}{2} g_m R_D \tag{9.34}$$

Alternatively, if the output is taken differentially, the gain becomes

$$A_d \equiv \frac{v_{od}}{v_{id}} = \frac{v_{o2} - v_{o1}}{v_{id}} = g_m R_D \tag{9.35}$$

Thus another advantage of taking the output differentially is an increase in gain by a factor of 2 (6 dB). It should be noted, however, that although differential outputs are preferred, a single-ended output is needed in some applications. We will have more to say about this later.

An alternative and useful way of viewing the operation of the differential pair in response to a differential input signal v_{id} is illustrated in Fig. 9.9. Here we are making use of the fact that the resistance between gate and source of a MOSFET, looking into the source, is $1/g_m$. As a result, between G_1 and G_2 we have a total resistance, in the source circuit, of $2/g_m$. It follows that we can obtain the current i_d simply by dividing v_{id} by $2/g_m$, as indicated in the figure.

The Differential Half-Circuit When a symmetrical differential amplifier is fed with a differential signal in a balanced manner, as in the case in Fig. 9.8, the performance can be determined by considering only half the circuit. The equivalent differential half-circuit is shown in Fig. 9.10. It has a grounded source, a result of the virtual ground that appears on the common sources' terminal of the MOSFETs in the differential pair. Note that Q_1 is operating at a drain bias current of $(I/2)$ and an overdrive voltage V_{OV}.

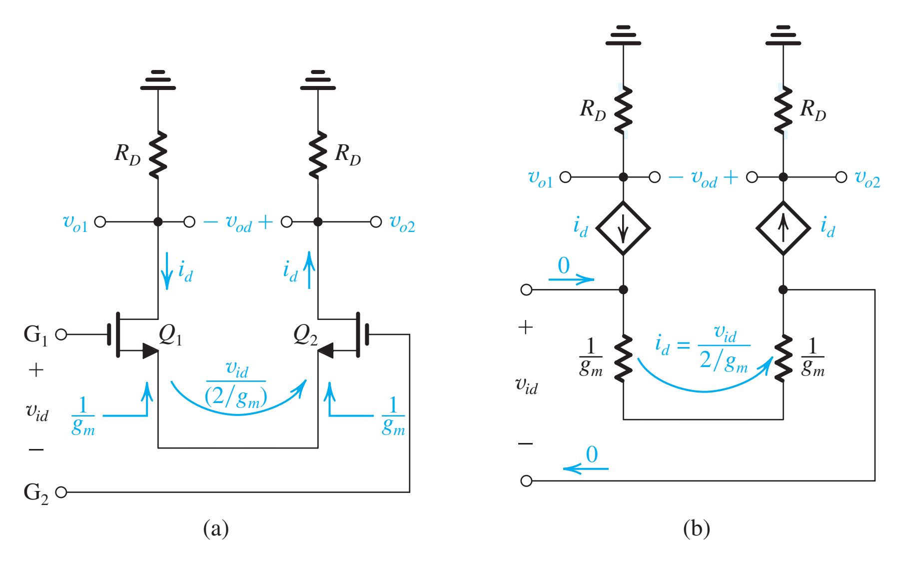


Figure 9.9 An alternative view of the small-signal differential operation of the MOS differential pair: (**a**) analysis done directly on the circuit; (**b**) analysis using equivalent-circuit models.

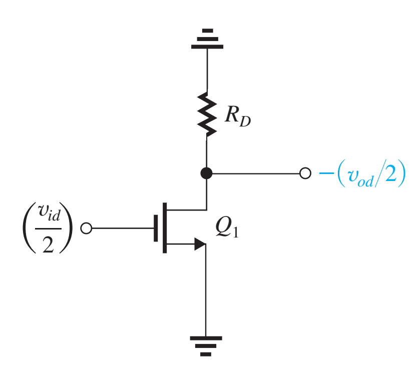
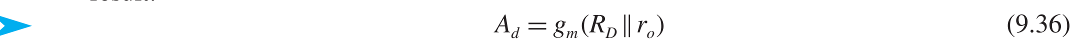


Figure 9.10 The equivalent differential half-circuit of the differential amplifier of Fig. 9.8. Here Q_1 is biased at $I/2$ and is operating at V_{OV}. This circuit can be used to determine the differential voltage gain of the differential amplifier $A_d = v_{od}/v_{id}$.

The differential gain A_d can be determined directly from the half-circuit. For instance, if we wish to take r_o of Q_1 and Q_2 into account, we can use the half-circuit with the following result:



$$A_d = g_m(R_D \| r_o) \tag{9.36}$$

More significantly, the frequency response of the differential gain can be determined by analyzing the half-circuit, as we shall do in Chapter 10.

Example 9.2

Give the differential half-circuit of the differential amplifier shown in Fig. 9.11(a). Assume that Q_1 and Q_2 are perfectly matched. Neglecting r_o, determine the differential voltage gain $A_d \equiv v_{od}/v_{id}$.

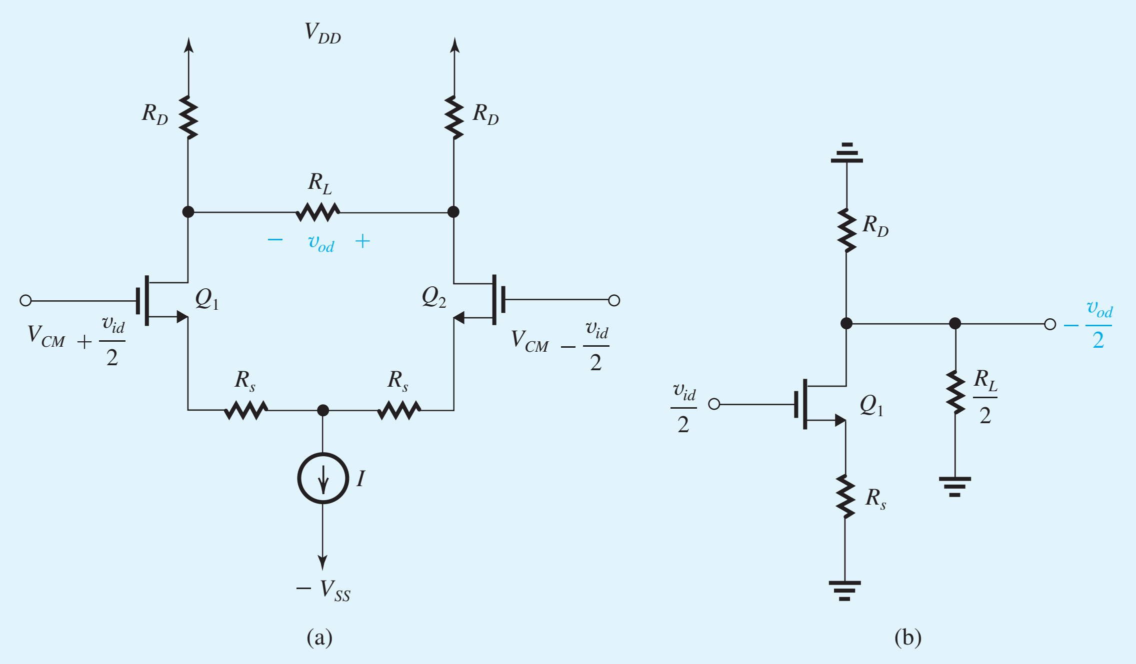


Figure 9.11 (**a**) Differential amplifier for Example 9.2. (**b**) Differential half-circuit.

Solution

Since the circuit is symmetrical and is fed with v_{id} in a balanced manner, the differential half-circuit will be as shown in Fig. 9.11(b). Observe that because the line of symmetry passes through the middle of R_L, the half-circuit has a resistance $R_L/2$ connected between drain and ground. Also note that the virtual ground appears on the node between the two resistances R_s. As a result, the half-circuit has a source-degeneration resistance R_s.

Now, neglecting r_o of the half-circuit transistor Q_1, we can obtain the gain as the ratio of the total resistance in the drain to the total resistance in the source as

$$\frac{-v_{od}/2}{v_{id}/2} = -\frac{R_D \| (R_L/2)}{1/g_m + R_s}$$

with the result that

$$A_d \equiv \frac{v_{od}}{v_{id}} = \frac{R_D \| (R_L/2)}{1/g_m + R_s} \tag{9.37}$$

EXERCISE

9.4 A MOS differential amplifier is operated at a total current of 0.8 mA, using transistors with a W/L ratio of 100, $\mu_n C_{ox} = 0.2\ \text{mA/V}^2$, $V_A = 20$ V, and $R_D = 5\ \text{k}\Omega$. Find V_{OV}, g_m, r_o, and A_d.
Ans. 0.2 V; 4 mA/V; 50 kΩ; 18.2 V/V

9.1.5 The Differential Amplifier with Current-Source Loads

To obtain higher gain, the passive resistances R_D can be replaced with current sources, as shown in Fig. 9.12(a). Here the current sources are realized with PMOS transistors Q_3 and Q_4, and V_G is a dc bias voltage that ensures that Q_3 and Q_4 each conducts a current equal to $I/2$. The differential voltage gain A_d can be found from the differential half-circuit shown in Fig. 9.12(b) as

$$A_d \equiv \frac{v_{od}}{v_{id}} = g_{m1}(r_{o1} \| r_{o3})$$

EXERCISE

9.5 The differential amplifier of Fig. 9.12(a) is fabricated in a 0.18-μm CMOS technology for which $\mu_n C_{ox} = 4\mu_p C_{ox} = 400\ \mu\text{A/V}^2$, $|V_t| = 0.5$ V, and $|V_A'| = 10$ V/μm. If the bias current $I = 200\ \mu$A and all transistors have a channel length twice the minimum and are operating at $|V_{OV}| = 0.2$ V, find W/L for each of Q_1, Q_2, Q_3, and Q_4, and determine the differential voltage gain A_d.
Ans. $(W/L)_{1,2} = 12.5$; $(W/L)_{3,4} = 50$; $A_d = 18$ V/V

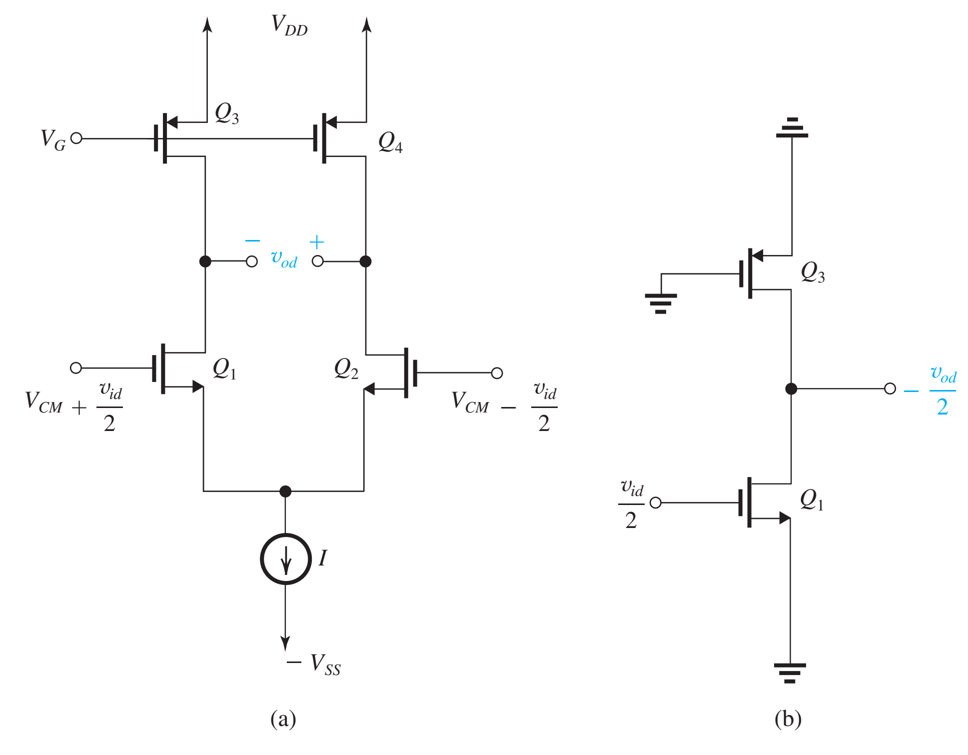


Figure 9.12 (**a**) Differential amplifier with current-source loads formed by Q_3 and Q_4. (**b**) Differential half-circuit of the amplifier in (**a**).

THE LONG-TAILED PAIR:

This idea using vacuum tubes was first documented by B. C. P. Matthews in 1934 in the *Proceedings of the Physical Society*, and was further developed by others in the late 1930s. The topology is simply that of a differential pair, where the term "long-tailed" refers to the biasing current source, which originally used a large-valued (hence long) resistor. Interestingly enough, the first application in measuring biological potentials in an electrically noisy environment continues to be an important one in modern medical instruments that utilize MOS devices.

9.1.6 Cascode Differential Amplifier

The gain of the differential amplifier can be increased by utilizing the cascode configuration studied in Section 8.5. Figure 9.13(a) shows a CMOS differential amplifier with cascoding applied to the amplifying transistors Q_1 and Q_2 via transistors Q_3 and Q_4, and to the current-source transistors Q_7 and Q_8 via transistors Q_5 and Q_6. The differential voltage gain can be found from the differential half-circuit shown in Fig. 9.13(b) as

$$A_d \equiv \frac{v_{od}}{v_{id}} = g_{m1}\left(R_{on} \| R_{op}\right) \tag{9.38}$$

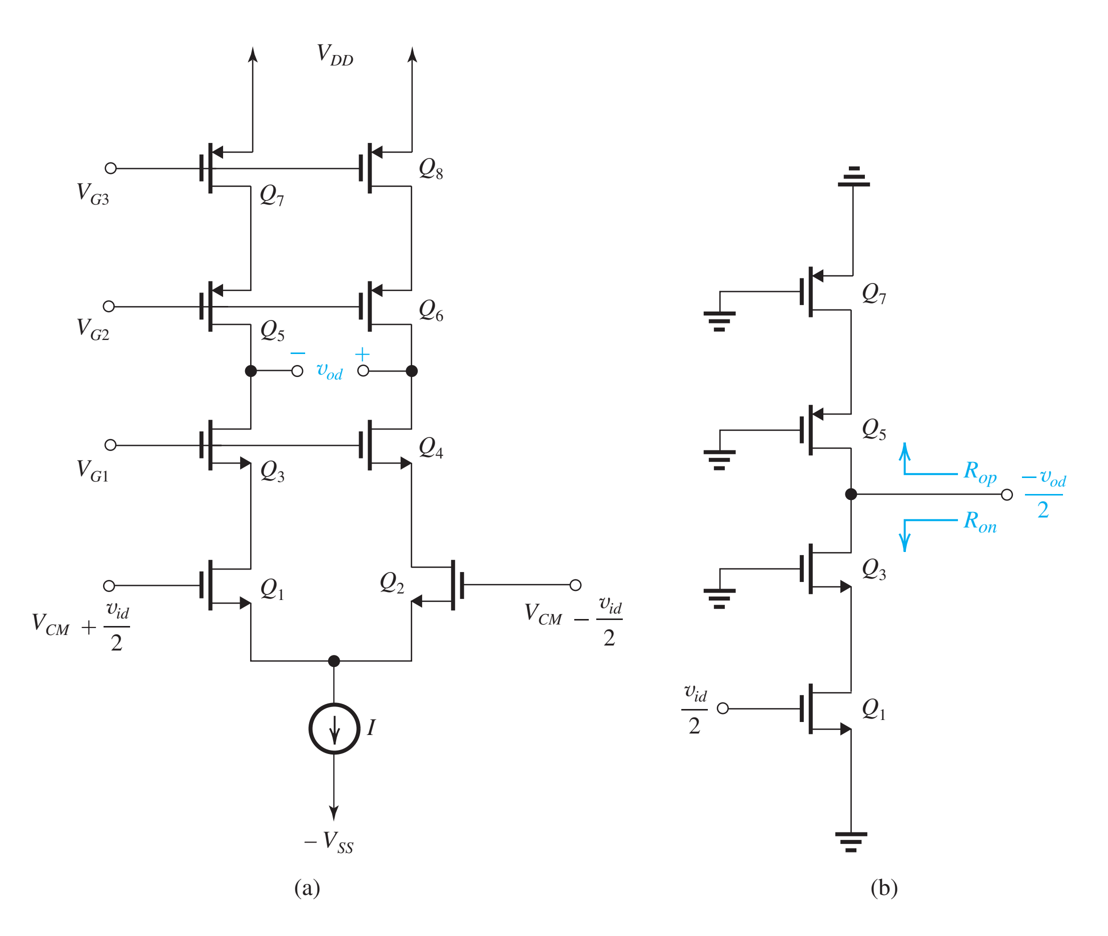


Figure 9.13 (**a**) Cascode differential amplifier; and (**b**) its differential half-circuit.

where

$$R_{on} = (g_{m3}r_{o3})r_{o1} \tag{9.39}$$

and,

$$R_{op} = (g_{m5}r_{o5})r_{o7} \tag{9.40}$$

EXERCISE

9.6 The CMOS cascode differential amplifier of Fig. 9.13(a) is fabricated in a 0.18-μm technology for which $\mu_n C_{ox} = 4\mu_p C_{ox} = 400\ \mu\text{A/V}^2$, $|V_t| = 0.5$ V, and $|V_A'| = 10$ V/μm. If the bias current $I = 200\ \mu$A, and all transistors have a channel length twice the minimum and are operating at $|V_{OV}| = 0.2$ V, find W/L for each of Q_1 to Q_8, and determine the differential voltage gain A_d.
Ans. $(W/L)_{1,2,3,4} = 12.5$; $(W/L)_{5,6,7,8} = 50$; $A_d = 648$ V/V

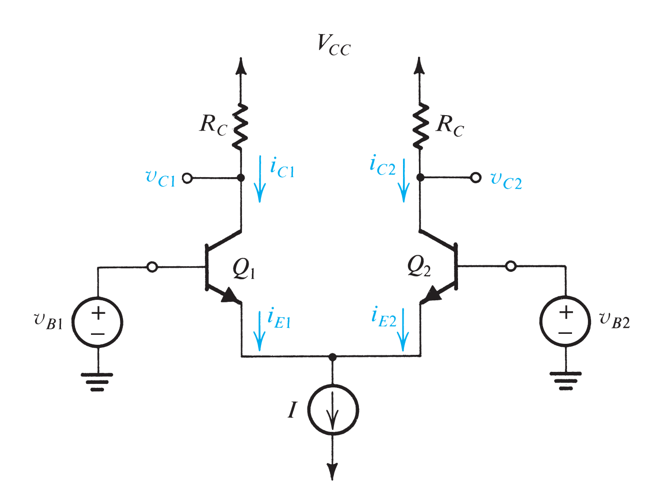


Figure 9.14 The basic BJT differential-pair configuration.

9.2 The BJT Differential Pair

Figure 9.14 shows the basic BJT differential-pair configuration. It is very similar to the MOSFET circuit and consists of two matched transistors, Q_1 and Q_2, whose emitters are joined together and biased by a constant-current source I. The latter is usually implemented by a transistor circuit of the type studied in Sections 8.2 and 8.6. Although each collector is shown connected to the positive supply voltage V_{CC} through a resistance R_C, this connection is not essential to the operation of the differential pair—that is, in some applications the two collectors may be connected to current sources rather than resistive loads. It is essential, though, that the collector circuits be such that Q_1 and Q_2 never enter saturation.

9.2.1 Basic Operation

To see how the BJT differential pair works, consider first the case of a common-mode voltage V_{CM} applied to the two input terminals. That is, as shown in Fig. 9.15(a), $v_{B1} = v_{B2} = V_{CM}$. Since Q_1 and Q_2 are matched, and assuming an ideal bias current source I with infinite output resistance, it follows that the current I will remain constant and, from symmetry, that I will divide equally between the two devices. Thus $i_{E1} = i_{E2} = I/2$, and the voltage at the emitters will be $V_{CM} - V_{BE}$, where V_{BE} is the base–emitter voltage [assumed in Fig 9.15(a) to be approximately 0.7 V] corresponding to an emitter current of $I/2$. The voltage at each collector will be $V_{CC} - \frac{1}{2}\alpha I R_C$, and the difference in voltage between the two collectors will be zero.

Now let us vary the value of the common-mode input voltage V_{CM}. Obviously, as long as Q_1 and Q_2 remain in the active region, and the current source I has sufficient voltage across it to operate properly, the current I will still divide equally between Q_1 and Q_2, and the voltages at the collectors will not change. Thus the differential pair does not respond to (i.e., it *rejects*) changes in the common-mode input voltage.

As another experiment, let the voltage v_{B2} be set to a constant value, say, zero (by grounding B_2), and let $v_{B1} = +1$ V [see Fig. 9.15(b)]. With a bit of reasoning it can be seen that Q_1 will be on and conducting all of the current I and that Q_2 will be off. For Q_1 to be on (with

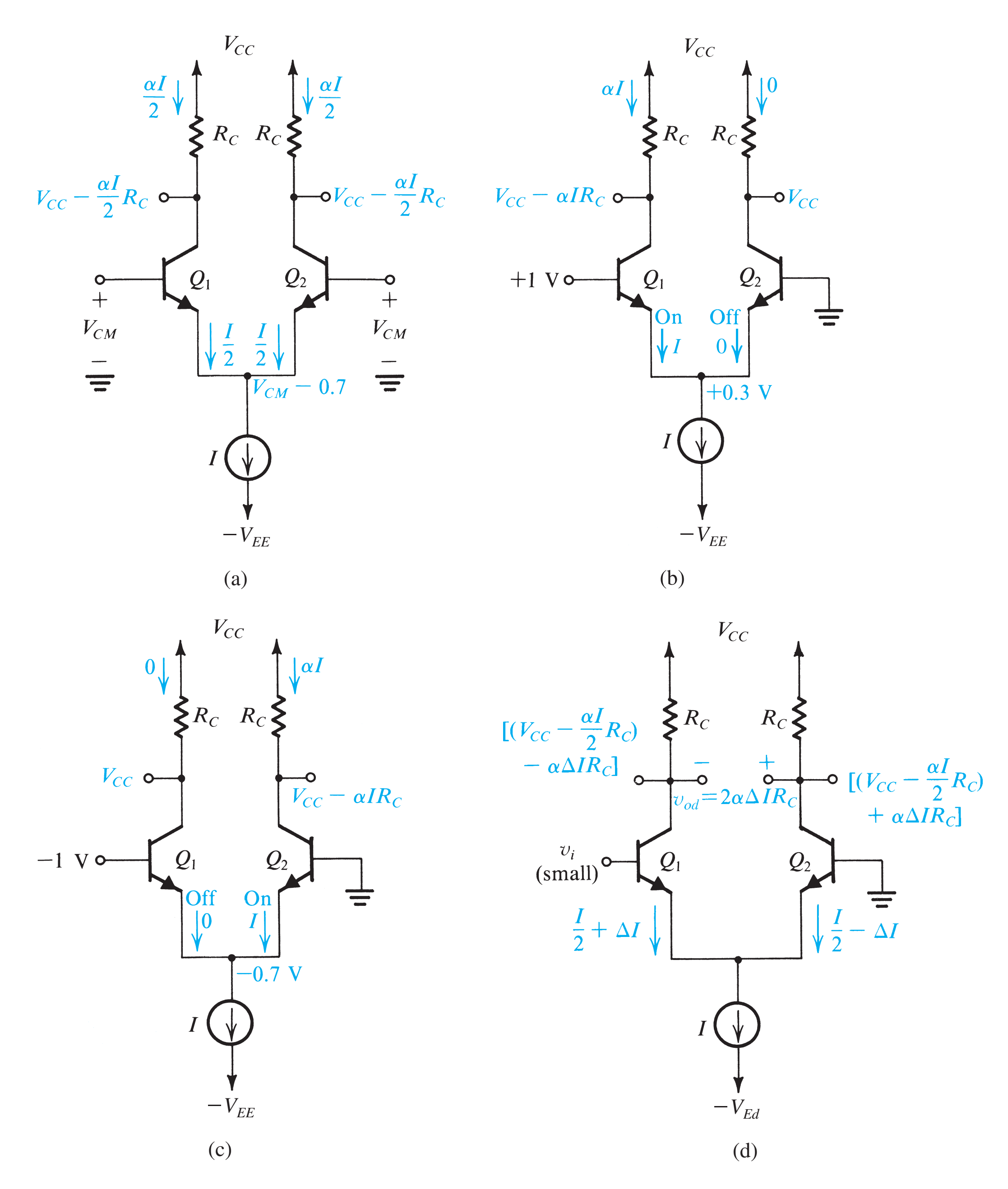


Figure 9.15 Different modes of operation of the BJT differential pair: **(a)** the differential pair with a common-mode input voltage V_{CM}; **(b)** the differential pair with a "large" differential input signal; **(c)** the differential pair with a large differential input signal of polarity opposite to that in **(b)**; **(d)** the differential pair with a small differential input signal v_i. Note that we have assumed the bias current source I to be ideal (i.e., it has an infinite output resistance) and thus I remains constant with the change in the voltage across it.

$V_{BE1} = 0.7$ V), the emitter has to be at approximately $+0.3$ V, which keeps the EBJ of Q_2 reverse-biased. The collector voltages will be $v_{C1} = V_{CC} - \alpha I R_C$ and $v_{C2} = V_{CC}$.

Let us now change v_{B1} to –1 V [Fig. 9.15(c)]. Again with some reasoning it can be seen that Q_1 will turn off, and Q_2 will carry all the current I. The common emitter will be at –0.7 V,

which means that the EBJ of Q_1 will be reverse biased by 0.3 V. The collector voltages will be $v_{C1} = V_{CC}$ and $v_{C2} = V_{CC} - \alpha IR_C$.

From the foregoing, we see that the differential pair certainly responds to difference-mode (or differential) signals. In fact, with relatively small difference voltages we are able to steer the entire bias current from one side of the pair to the other. This current-steering property of the differential pair allows it to be used in logic circuits, as will be demonstrated in Chapter 15.

To use the BJT differential pair as a linear amplifier, we apply a very small differential signal (a few millivolts), which will result in one of the transistors conducting a current of $I/2 + \Delta I$; the current in the other transistor will be $I/2 - \Delta I$, with ΔI being proportional to the difference input voltage (see Fig. 9.15(d)). The output voltage taken between the two collectors will be $2\alpha \Delta IR_C$, which is proportional to the differential input signal v_i. The small-signal operation of the differential pair will be studied shortly.

EXERCISE

9.7 Find v_E, v_{C1}, and v_{C2} in the circuit of Fig. E9.7. Assume that $|v_{BE}|$ of a conducting transistor is approximately 0.7 V and that $\alpha \simeq 1$.

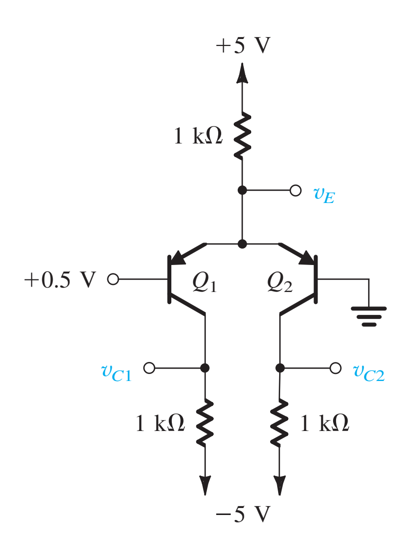


Figure E9.7

Ans. +0.7 V; −5 V; −0.7 V

9.2.2 Input Common-Mode Range

Refer to the circuit in Fig. 9.15(a). The allowable range of V_{CM} is determined at the upper end by Q_1 and Q_2 leaving the active mode and entering saturation. Thus

$$V_{CM\max} \simeq V_C + 0.4 = V_{CC} - \alpha \frac{I}{2} R_C + 0.4 \tag{9.41}$$

The lower end of the V_{CM} range is determined by the need to provide a certain minimum voltage V_{CS} across the current source I to ensure its proper operation. Thus,

$$V_{CM\min} = -V_{EE} + V_{CS} + V_{BE} \tag{9.42}$$

EXERCISE

9.8 Determine the input common-mode range for a bipolar differential amplifier operating from ±2.5-V power supplies and biased with a simple current source that delivers a constant current of 0.4 mA and requires a minimum of 0.3 V for its proper operation. The collector resistances $R_C = 5\ \text{k}\Omega$.
Ans. −1.5 V to +1.9 V

9.2.3 Large-Signal Operation

We now present a general analysis of the BJT differential pair of Fig. 9.14. If we denote the voltage at the common emitter by v_E and neglect the Early effect, the exponential relationship applied to each of the two transistors may be written

$$i_{E1} = \frac{I_S}{\alpha} e^{(v_{B1} - v_E)/V_T} \tag{9.43}$$

$$i_{E2} = \frac{I_S}{\alpha} e^{(v_{B2} - v_E)/V_T} \tag{9.44}$$

These two equations can be combined to obtain

$$\frac{i_{E1}}{i_{E2}} = e^{(v_{B1} - v_{B2})/V_T}$$

which can be manipulated to yield

$$\frac{i_{E1}}{i_{E1} + i_{E2}} = \frac{1}{1 + e^{(v_{B2} - v_{B1})/V_T}} \tag{9.45}$$

$$\frac{i_{E2}}{i_{E1} + i_{E2}} = \frac{1}{1 + e^{(v_{B1} - v_{B2})/V_T}} \tag{9.46}$$

The circuit imposes the additional constraint

$$i_{E1} + i_{E2} = I \tag{9.47}$$

Using Eq. (9.47) together with Eqs. (9.45) and (9.46) and substituting $v_{B1} - v_{B2} = v_{id}$ gives

$$i_{E1} = \frac{I}{1 + e^{-v_{id}/V_T}} \tag{9.48}$$

$$i_{E2} = \frac{I}{1 + e^{v_{id}/V_T}} \tag{9.49}$$

The collector currents i_{C1} and i_{C2} can be obtained simply by multiplying the emitter currents in Eqs. (9.48) and (9.49) by α, which is normally very close to unity.

The fundamental operation of the differential amplifier is illustrated by Eqs. (9.48) and (9.49). First, note that the amplifier responds only to the difference voltage v_{id}. That is, if $v_{B1} = v_{B2} = V_{CM}$, the current I divides equally between the two transistors irrespective of the value of the common-mode voltage V_{CM}. This is the essence of differential-amplifier operation, which also gives rise to its name.

Another important observation is that a relatively small difference voltage v_{id} will cause the current I to flow almost entirely in one of the two transistors. Figure 9.16 shows a plot of the two collector currents (assuming $\alpha \simeq 1$) as a function of the differential input signal. This is a normalized plot that can be used universally. Observe that a difference voltage of about $4V_T$ ($\simeq$100 mV) is sufficient to switch the current almost entirely to one side of the BJT pair. Note that this is much smaller than the corresponding voltage for the MOS pair, $\sqrt{2}\, V_{OV}$. The fact that such a small signal can switch the current from one side of the BJT differential pair to the other means that the BJT differential pair can be used as a fast current switch (Chapter 15).

The nonlinear transfer characteristics of the differential pair, shown in Fig. 9.16, will not be utilized any further in this chapter. Rather, in the following we shall be interested specifically in the application of the differential pair as a small-signal amplifier. For this purpose, the difference input signal is limited to less than about $V_T/2$ in order that we may operate on a linear segment of the characteristics around the midpoint x (in Fig. 9.16).

Before leaving the large-signal operation of the differential BJT pair, we wish to point out an effective technique frequently employed to extend the linear range of operation. It consists of including two equal resistances R_e in series with the emitters of Q_1 and Q_2, as

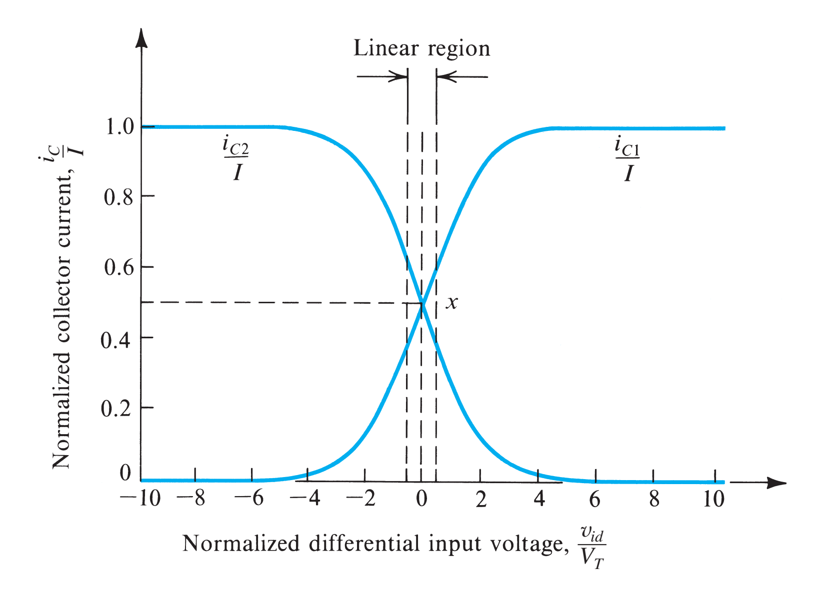


Figure 9.16 Transfer characteristics of the BJT differential pair of Fig. 9.14 assuming $\alpha \simeq 1$.

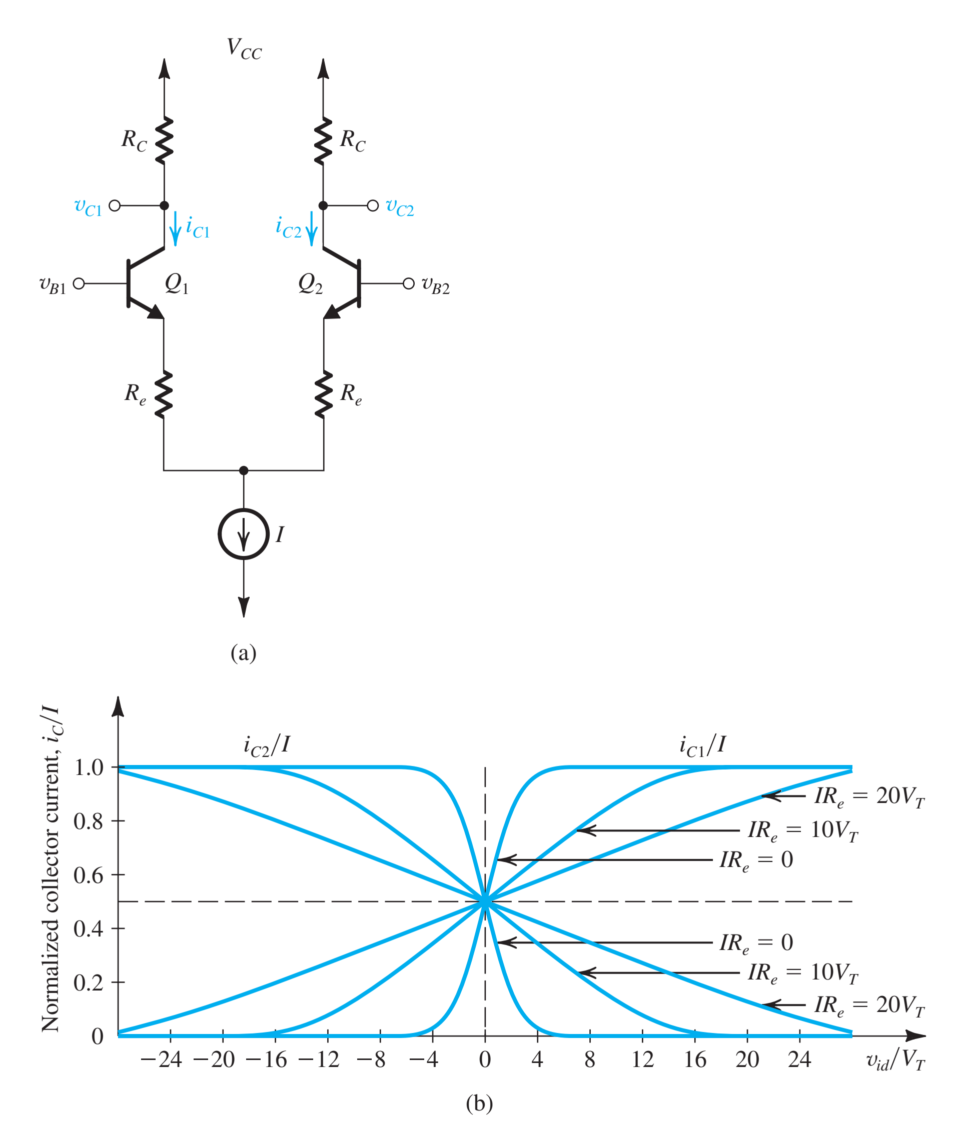


Figure 9.17 The transfer characteristics of the BJT differential pair **(a)** can be linearized **(b)** (i.e., the linear range of operation can be extended) by including resistances in the emitters.

shown in Fig. 9.17(a). The resulting transfer characteristics for three different values of R_e are sketched in Fig. 9.17(b). Observe that expansion of the linear range is obtained at the expense of reduced G_m (which is the slope of the transfer curve at $v_{id} = 0$) and hence reduced gain. This result should come as no surprise; R_e here is performing in exactly the same way as the emitter resistance R_e does in the CE amplifier with emitter degeneration (see Section 7.3.4). Finally, we also note that this linearization technique is in effect the bipolar counterpart of the technique employed for the MOS differential pair (Fig. 9.7). In the latter case, however, V_{OV} was varied by changing the transistors' W/L ratio, a design tool with no counterpart in the BJT.

EXERCISE

9.9 For the BJT differential pair of Fig. 9.14, find the value of input differential signal that is sufficient to cause $i_{E1} = 0.99I$.
Ans. 115 mV

9.2.4 Small-Signal Operation

In this section we study the application of the BJT differential pair in small-signal amplification. Figure 9.18 shows the BJT differential pair with a difference voltage signal v_{id} applied between the two bases. Implied is that the dc level at the input—that is, the common-mode input voltage—has been somehow established. For instance, one of the two input terminals can be grounded and v_{id} applied to the other input terminal. Alternatively, the differential amplifier may be fed from the output of another differential amplifier. In the latter case, the voltage at one of the input terminals will be $V_{CM} + v_{id}/2$ while that at the other input terminal will be $V_{CM} - v_{id}/2$.

The Collector Currents When v_{id} Is Applied For the circuit of Fig. 9.18, we may use Eqs. (9.48) and (9.49) to write

$$i_{C1} = \frac{\alpha I}{1 + e^{-v_{id}/V_T}} \tag{9.50}$$

$$i_{C2} = \frac{\alpha I}{1 + e^{v_{id}/V_T}} \tag{9.51}$$

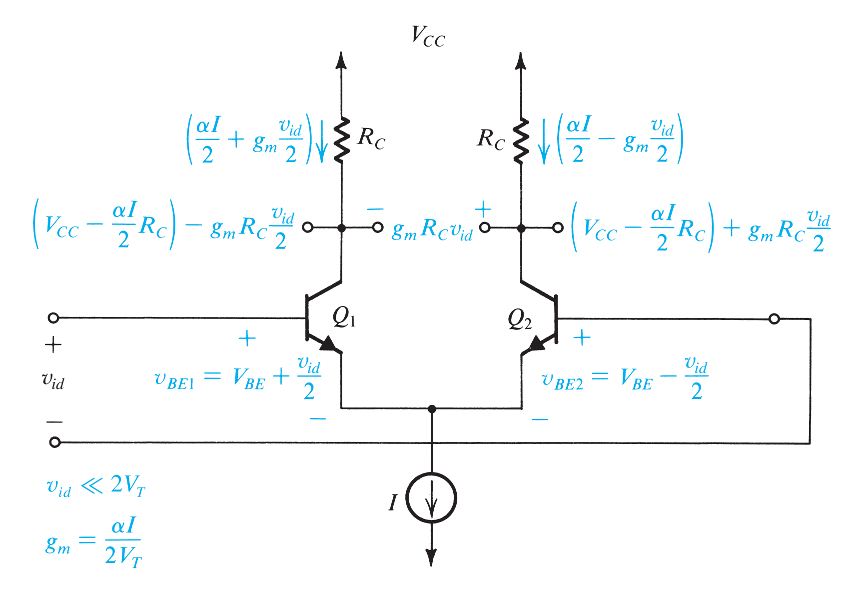


Figure 9.18 The currents and voltages in the differential amplifier when a small differential input signal v_{id} is applied.

Multiplying the numerator and the denominator of the right-hand side of Eq. (9.50) by $e^{v_{id}/2V_T}$ gives

$$i_{C1} = \frac{\alpha I e^{v_{id}/2V_T}}{e^{v_{id}/2V_T} + e^{-v_{id}/2V_T}}$$

Assume that $v_{id} \ll 2V_T$. We may thus expand the exponential $e^{\pm v_{id}/2V_T}$ in a series and retain only the first two terms:

$$i_{C1} \simeq \frac{\alpha I(1 + v_{id}/2V_T)}{1 + v_{id}/2V_T + 1 - v_{id}/2V_T}$$

Thus

$$i_{C1} = \frac{\alpha I}{2} + \frac{\alpha I}{2V_T}\frac{v_{id}}{2} \tag{9.52}$$

Similar manipulations can be applied to Eq. (9.51) to obtain

$$i_{C2} = \frac{\alpha I}{2} - \frac{\alpha I}{2V_T}\frac{v_{id}}{2} \tag{9.53}$$

Equations (9.52) and (9.53) tell us that when $v_{id} = 0$, the bias current I divides equally between the two transistors of the pair. Thus each transistor is biased at an emitter current of $I/2$. When a "small-signal" v_{id} is applied differentially (i.e., between the two bases), the collector current of Q_1 increases by an increment i_c and that of Q_2 decreases by an equal amount. This ensures that the sum of the total currents in Q_1 and Q_2 remains constant, as constrained by the current-source bias. The incremental (or signal) current component i_c is given by

$$i_c = \frac{\alpha I}{2V_T}\frac{v_{id}}{2} \tag{9.54}$$

Equation (9.54) has an easy interpretation. First, note from the symmetry of the circuit (Fig. 9.18) that the differential signal v_{id} should divide equally between the base–emitter junctions of the two transistors. Thus the total base–emitter voltages will be

$$v_{BE}\big|_{Q1} = V_{BE} + \frac{v_{id}}{2}$$

$$v_{BE}\big|_{Q2} = V_{BE} - \frac{v_{id}}{2}$$

where V_{BE} is the dc BE voltage corresponding to an emitter current of $I/2$. Therefore, the collector current of Q_1 will increase by $g_m v_{id}/2$ and the collector current of Q_2 will decrease by $g_m v_{id}/2$. Here g_m denotes the transconductance of Q_1 and of Q_2, which are equal and given by

$$g_m = \frac{I_C}{V_T} = \frac{\alpha I/2}{V_T} \tag{9.55}$$

Thus Eq. (9.54) simply states that $i_c = g_m v_{id}/2$.

An Alternative Viewpoint There is an extremely useful alternative interpretation of the results above. Assume the current source I to be ideal. Its incremental resistance then will be infinite. Thus the voltage v_{id} appears across a total resistance of $2r_e$, where

$$r_e = \frac{V_T}{I_E} = \frac{V_T}{I/2} \tag{9.56}$$

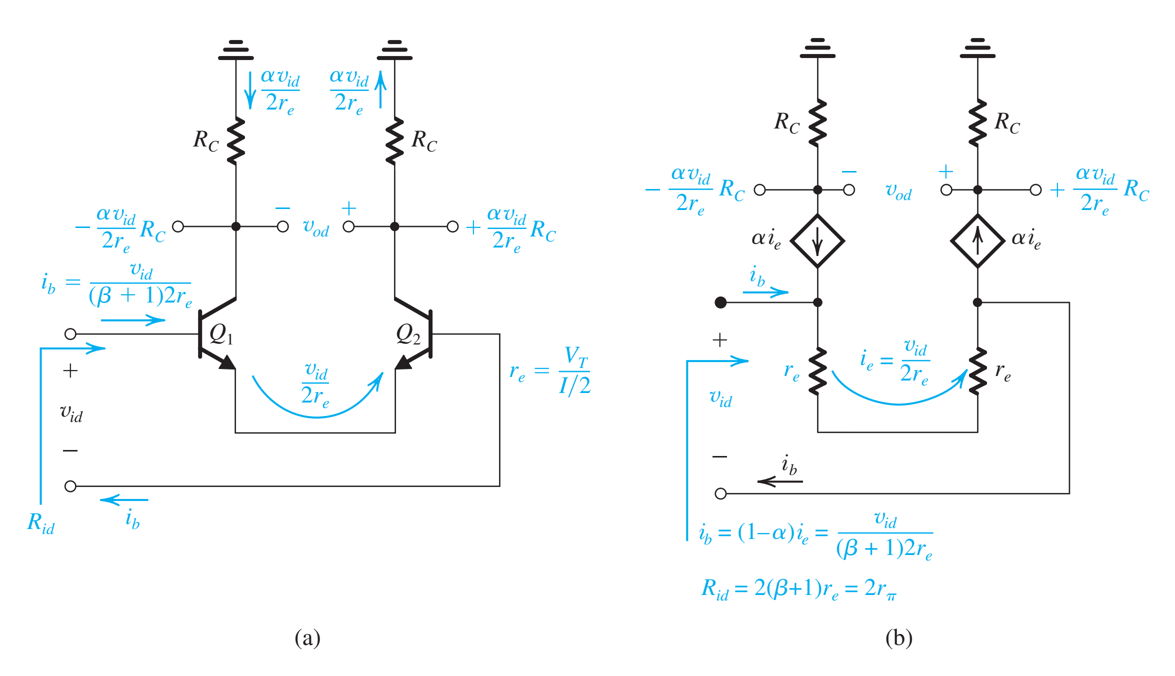


Figure 9.19 A simple technique for determining the signal currents in a differential amplifier excited by a differential voltage signal v_{id}; dc quantities are not shown. While Fig. 9.19(a) utilizes the BJT T model implicitly, the T model of both BJTs are shown explicitly in Fig. 9.19(b).

Correspondingly there will be a signal current i_e, as illustrated in Fig. 9.19, given by

$$i_e = \frac{v_{id}}{2r_e} \tag{9.57}$$

Thus the collector of Q_1 will exhibit a current increment i_c and the collector of Q_2 will exhibit a current decrement i_c:

$$i_c = \alpha i_e = \frac{\alpha v_{id}}{2r_e} = g_m \frac{v_{id}}{2} \tag{9.58}$$

Note that in Fig. 9.19(a) we have shown signal quantities only. It is implied, of course, that each transistor is biased at an emitter current of $I/2$. For greater emphasis, we show in Fig. 9.20(b) the equivalent circuit obtained by replacing each BJT with its T model.

This method of analysis is particularly useful when resistances are included in the emitters, as shown in Fig. 9.20. For this circuit we have

$$i_e = \frac{v_{id}}{2r_e + 2R_e} \tag{9.59}$$

Input Differential Resistance Unlike the MOS differential amplifier, which has an infinite input resistance, the bipolar differential pair exhibits a finite input resistance, a result of the finite β of the BJT.

The input differential resistance is the resistance seen between the two bases; that is, it is the resistance seen by the differential input signal v_{id}. For the differential amplifier in Figs. 9.18 and 9.19 it can be seen that the base current of Q_1 shows an increment i_b and the base current of Q_2 shows an equal decrement,

$$i_b = \frac{i_e}{\beta+1} = \frac{v_{id}/2r_e}{\beta+1} \tag{9.60}$$

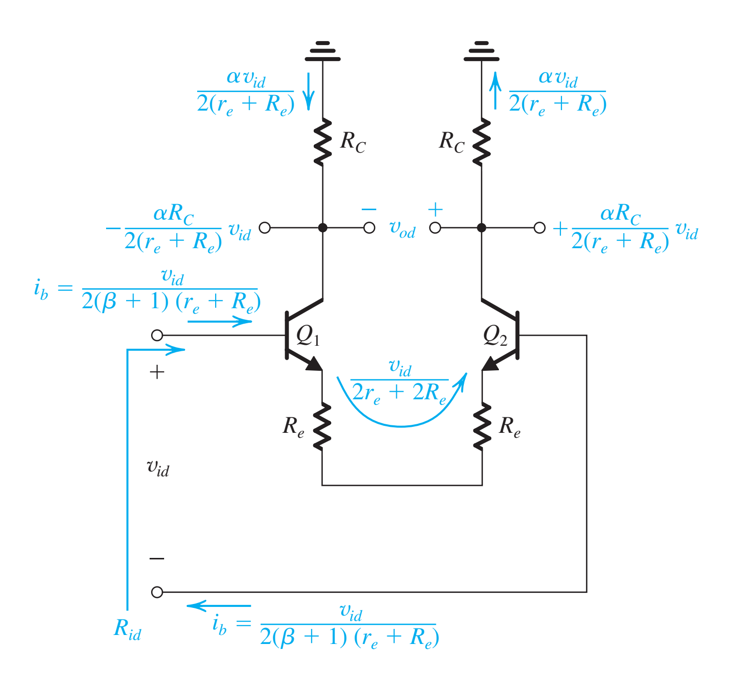


Figure 9.20 A differential amplifier with emitter resistances. Only signal quantities are shown (in color).

Thus the differential input resistance R_{id} is given by

$$R_{id} \equiv \frac{v_{id}}{i_b} = (\beta + 1)2r_e = 2r_\pi \tag{9.61}$$

This result is just a restatement of the familiar resistance-reflection rule: namely, *the resistance seen between the two bases is equal to the total resistance in the emitter circuit multiplied by* $(\beta + 1)$. We can employ this rule to find the input differential resistance for the circuit in Fig. 9.20 as

$$R_{id} = (\beta + 1)(2r_e + 2R_e) \tag{9.62}$$

Differential Voltage Gain We have established that for small difference input voltages ($v_{id} \ll 2\,V_T$; i.e., v_{id} smaller than about 20 mV), the collector currents are given by

$$i_{C1} = I_C + g_m \frac{v_{id}}{2} \tag{9.63}$$

$$i_{C2} = I_C - g_m \frac{v_{id}}{2} \tag{9.64}$$

where

$$I_C = \frac{\alpha I}{2} \tag{9.65}$$

Thus the total voltages at the collectors will be

$$v_{C1} = (V_{CC} - I_C R_C) - g_m R_C \frac{v_{id}}{2} \tag{9.66}$$

$$v_{C2} = (V_{CC} - I_C R_C) + g_m R_C \frac{v_{id}}{2} \tag{9.67}$$

The quantities in parentheses are simply the dc voltages at each of the two collectors.

As in the MOS case, the output voltage signal of a bipolar differential amplifier can be taken *differentially* (i.e., between the two collectors, $v_{od} = v_{c2} - v_{c1}$). The differential gain of the differential amplifier will be

$$A_d = \frac{v_{od}}{v_{id}} = g_m R_C \tag{9.68}$$

For the differential amplifier with resistances in the emitter leads (Fig. 9.20), the differential gain is given by

$$A_d = \frac{\alpha(2R_C)}{2r_e + 2R_e} \simeq \frac{R_C}{r_e + R_e} \tag{9.69}$$

This equation is a familiar one: It states that *the voltage gain is equal to the ratio of the total resistance in the collector circuit* ($2R_C$) *to the total resistance in the emitter circuit* ($2r_e + 2R_e$).

The Differential Half-Circuit As in the MOS case, the differential gain of the BJT differential amplifier can be obtained by considering its differential half-circuit. Figure 9.21(a) shows a differential amplifier fed by a differential signal v_{id} that is applied in a **complementary** (**push–pull** or **balanced**) manner. That is, while the base of Q_1 is raised by $v_{id}/2$, the base of Q_2 is lowered by $v_{id}/2$. We have also included the output resistance R_{EE} of the bias current source. From symmetry, it follows that the signal voltage at the emitters will be zero. Thus the circuit is equivalent to the two common-emitter amplifiers shown in Fig. 9.21(b), where each of the two transistors is biased at an emitter current of $I/2$. Note that the finite output resistance R_{EE} of the current source will have no effect on the operation. The equivalent circuit in Fig. 9.21(b) is valid for differential operation only.

In many applications the differential amplifier is not fed in a complementary fashion; rather, the input signal may be applied to one of the input terminals while the other terminal is

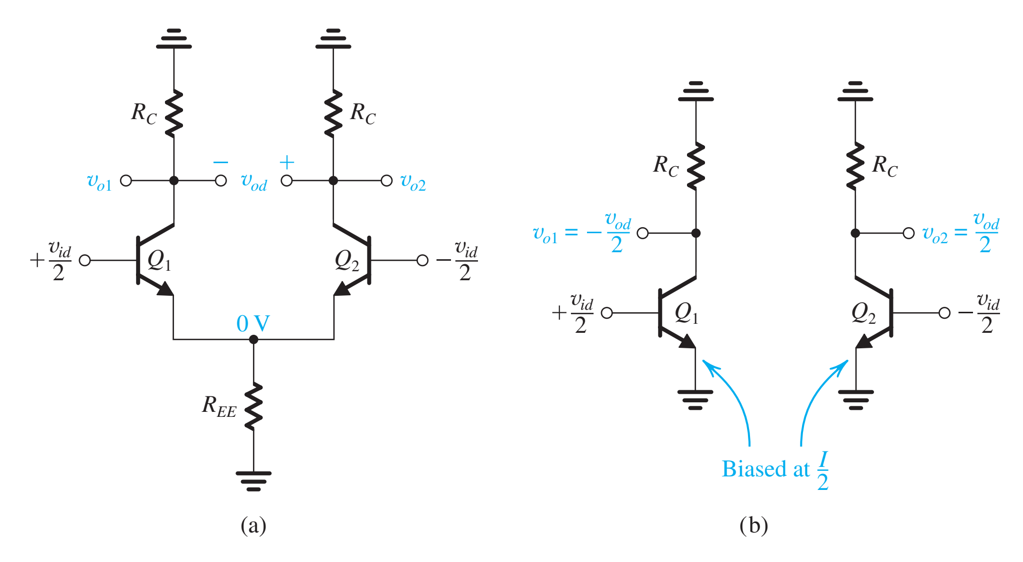


Figure 9.21 Equivalence of the BJT differential amplifier in **(a)** to the two common-emitter amplifiers in **(b).** This equivalence applies only for differential input signals. Either of the two common-emitter amplifiers in **(b)** can be used to find the differential gain, differential input resistance, frequency response, and so on, of the differential amplifier.

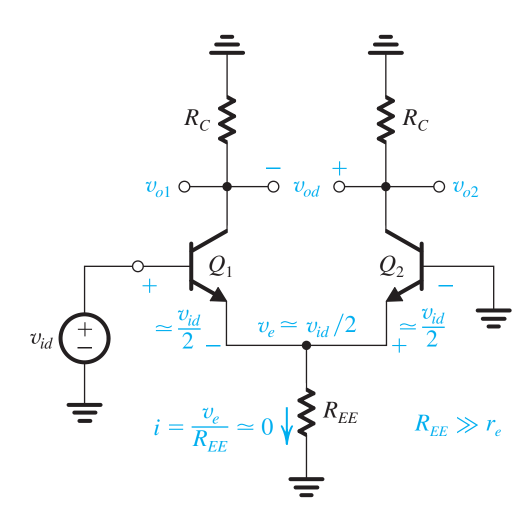


Figure 9.22 The differential amplifier fed in a single-ended fashion.

grounded, as shown in Fig. 9.22. In this case the signal voltage at the emitters will not be zero, and thus the resistance R_{EE} will have an effect on the operation. Nevertheless, if R_{EE} is large ($R_{EE} \gg r_e$), as is usually the case,[1] then v_{id} will still divide equally (approximately) between the two junctions, as shown in Fig. 9.22. Thus the operation of the differential amplifier in this case will be almost identical to that in the case of symmetric feed, and the common-emitter equivalence can still be employed.

Since in Fig. 9.21, $v_{o2} = -v_{o1} = v_{od}/2$, the two common-emitter transistors in Fig. 9.21(b) yield similar results about the performance of the differential amplifier. Thus only one is needed to analyze the differential small-signal operation of the differential amplifier, and it is known as the **differential half-circuit**. If we take the common-emitter transistor fed with $+v_{id}/2$ as the differential half-circuit and replace the transistor with its low-frequency, hybrid-π, equivalent-circuit model, the circuit in Fig. 9.23 results. In evaluating the model parameters r_π, g_m, and r_o, we must recall that the half-circuit is biased at $I/2$. The voltage gain of the differential amplifier is equal to the voltage gain of the half-circuit—that is, $v_{o1}/(v_{id}/2)$. Here, we note that including r_o will modify the gain expression in Eq. (9.68) to

$$A_d = g_m(R_C \parallel r_o) \tag{9.70}$$

The input differential resistance of the differential amplifier is twice that of the half-circuit—that is, $2r_\pi$. Finally, we note that the differential half-circuit of the amplifier of Fig. 9.20 is a common-emitter transistor with a resistance R_e in the emitter lead.

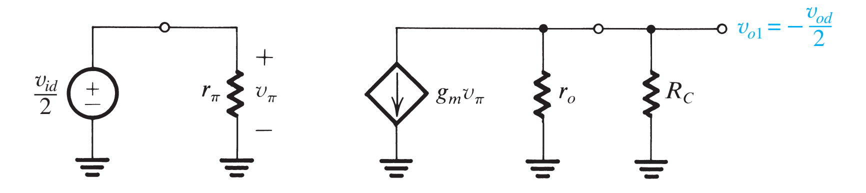


Figure 9.23 Equivalent-circuit model of the differential half-circuit formed by Q_1 in Fig. 9.22(b).

[1]Note that R_{EE} appears in parallel with the much smaller r_e of Q_2.

Example 9.3

The differential amplifier in Fig. 9.24 uses transistors with $\beta = 100$. Evaluate the following:

(a) The input differential resistance R_{id}.
(b) The overall differential voltage gain v_{od}/v_{sig} (neglect the effect of r_o).

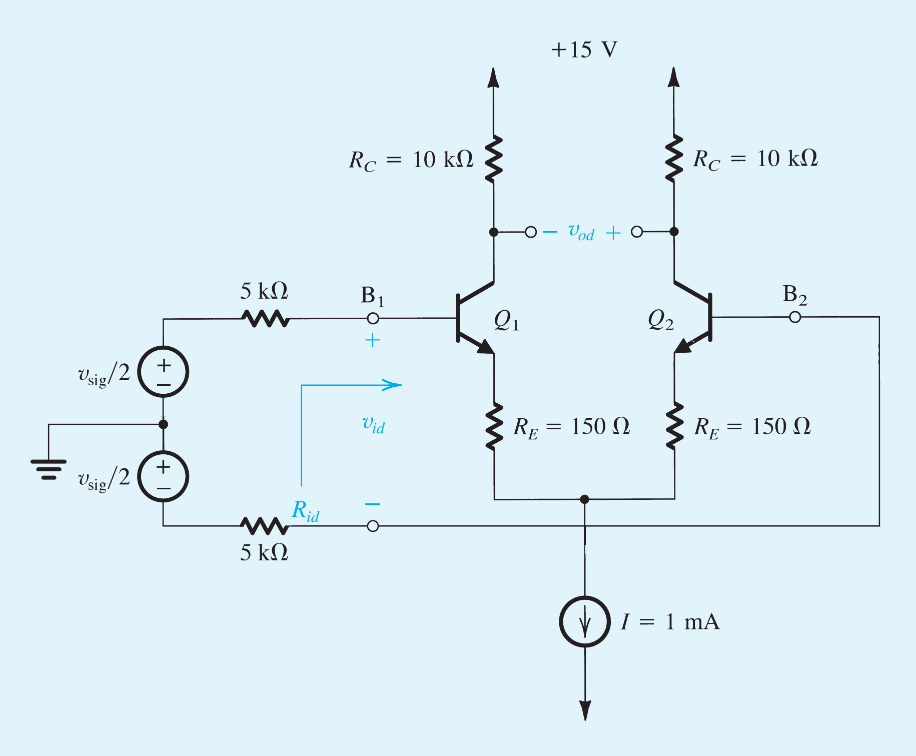


Figure 9.24 Circuit for Example 9.3.

Solution

(a) Each transistor is biased at an emitter current of 0.5 mA. Thus

$$r_{e1} = r_{e2} = \frac{V_T}{I_E} = \frac{25\text{ mV}}{0.5\text{ mA}} = 50\ \Omega$$

The input differential resistance can now be found as

$$R_{id} = 2(\beta + 1)\left(r_e + R_E\right)$$
$$= 2 \times 101 \times (50 + 150) \simeq 40\text{ k}\Omega$$

(b) The voltage gain from the signal source to the bases of Q_1 and Q_2 is

$$\frac{v_{id}}{v_{sig}} = \frac{R_{id}}{R_{sig} + R_{id}}$$
$$= \frac{40}{5+5+40} = 0.8 \text{ V/V}$$

The voltage gain from the bases to the output is

$$\frac{v_{od}}{v_{id}} \simeq \frac{\text{Total resistance in the collectors}}{\text{Total resistance in the emitters}}$$
$$= \frac{2R_C}{2(r_e + R_E)} = \frac{2 \times 10}{2(50+150) \times 10^{-3}} = 50 \text{ V/V}$$

The overall differential voltage gain can now be found as

$$A_d = \frac{v_{od}}{v_{sig}} = \frac{v_{id}}{v_{sig}} \frac{v_{od}}{v_{id}} = 0.8 \times 50 = 40 \text{ V/V}$$

EXERCISE

9.10 For the circuit in Fig. 9.18, let $I = 1$ mA, $V_{CC} = 15$ V, $R_C = 10\text{ k}\Omega$, with $\alpha = 1$, and let the input voltages be $v_{B1} = 5 + 0.005 \sin 2\pi \times 1000t$, volts, and $v_{B2} = 5 - 0.005 \sin 2\pi \times 1000t$, volts. (a) If the BJTs are specified to have v_{BE} of 0.7 V at a collector current of 1 mA, find the voltage at the emitters. (b) Find g_m for each of the two transistors. (c) Find i_C for each of the two transistors. (d) Find v_C for each of the two transistors. (e) Find the voltage between the two collectors. (f) Find the gain experienced by the 1000-Hz signal.
Ans. (a) 4.317 V; (b) 20 mA/V; (c) $i_{C1} = 0.5 + 0.1 \sin 2\pi \times 1000t$, mA and $i_{C2} = 0.5 - 0.1 \sin 2\pi \times 1000t$, mA; (d) $v_{C1} = 10 - 1 \sin 2\pi \times 1000t$, V and $v_{C2} = 10 + 1 \sin 2\pi \times 1000t$, V; (e) $v_{C2} - v_{C1} = 2 \sin 2\pi \times 1000t$, V; (f) 200 V/V

9.3 Common-Mode Rejection

Thus far, we have seen that the differential amplifier responds to a differential input signal and completely rejects a common-mode signal. This latter point was made very clearly at the outset of our discussion of differential amplifiers and was illustrated in Example 9.1, where we saw that changes in V_{CM} over a wide range resulted in no change in the voltage at either of the two drains. The same phenomenon was demonstrated for the BJT differential amplifier in Section 9.2.1. This highly desirable result is, however, a consequence of our assumption that the current source that supplies the bias current I is ideal. As we shall now show, if we

consider the more realistic situation of the current source having a finite output resistance, the common-mode gain will no longer be zero.

9.3.1 The MOS Case

Figure 9.25(a) shows a MOS differential amplifier biased with a current source having an output resistance R_{SS}. As before, the dc voltage at the input is defined by V_{CM}. Here, however, we also have an incremental signal v_{icm} applied to both input terminals. This common-mode input signal can represent an interference signal or noise that is picked up by both inputs and is clearly undesirable. Our objective now is to find how much of v_{icm} makes its way to the output of the amplifier.

Before we determine the common-mode gain of the amplifier, we wish to address the question of the effect of R_{SS} on the bias current of Q_1 and Q_2. That is, with v_{icm} set to zero, the bias current in each of Q_1 and Q_2 will no longer be $I/2$ but will be larger than $I/2$ by an amount determined by V_{CM} and R_{SS}. However, since R_{SS} is usually very large, this additional dc current in each of Q_1 and Q_2 is usually small and we shall neglect it, thus assuming that Q_1 and Q_2 continue to operate at a bias current of $I/2$. The reader might also be wondering about the effect of R_{SS} on the differential gain. The answer here is very simple: The virtual ground that develops on the common-source terminal when a differential input signal is applied results in a zero signal current through R_{SS}; hence R_{SS} has no effect on the value of A_d.

To determine the response of the differential amplifier to the common-mode input signal v_{icm}, consider the circuit in Fig. 9.25(b), where we have replaced each of V_{DD} and V_{SS} by a short circuit and I by an open circuit. The circuit is obviously symmetrical, and thus the two transistors will carry equal signal currents, denoted i. The value of i can be easily determined by replacing each of Q_1 and Q_2 with its T model and, for simplicity, neglecting r_o. The resulting equivalent circuit is shown in Fig. 9.25(c), from which we can write

$$v_{icm} = \frac{i}{g_m} + 2iR_{SS} \tag{9.71}$$

Thus,

$$i = \frac{v_{icm}}{1/g_m + 2R_{SS}} \tag{9.72}$$

The voltages at the drain of Q_1 and Q_2 can now be found as

$$v_{o1} = v_{o2} = -R_D i$$

resulting in

$$v_{o1} = v_{o2} = -\frac{R_D}{1/g_m + 2R_{SS}} v_{icm} \tag{9.73}$$

It follows that both v_{o1} and v_{o2} will be corrupted by the common-mode signal v_{icm} and will be given approximately by

$$\frac{v_{o1}}{v_{icm}} = \frac{v_{o2}}{v_{icm}} \simeq -\frac{R_D}{2R_{SS}} \tag{9.74}$$

where we have assumed that $2R_{SS} \gg 1/g_m$. Nevertheless, because $v_{o1} = v_{o2}$, the differential output voltage v_{od} will remain free of common-mode interference:

$$v_{od} = v_{o2} - v_{o1} = 0 \tag{9.75}$$

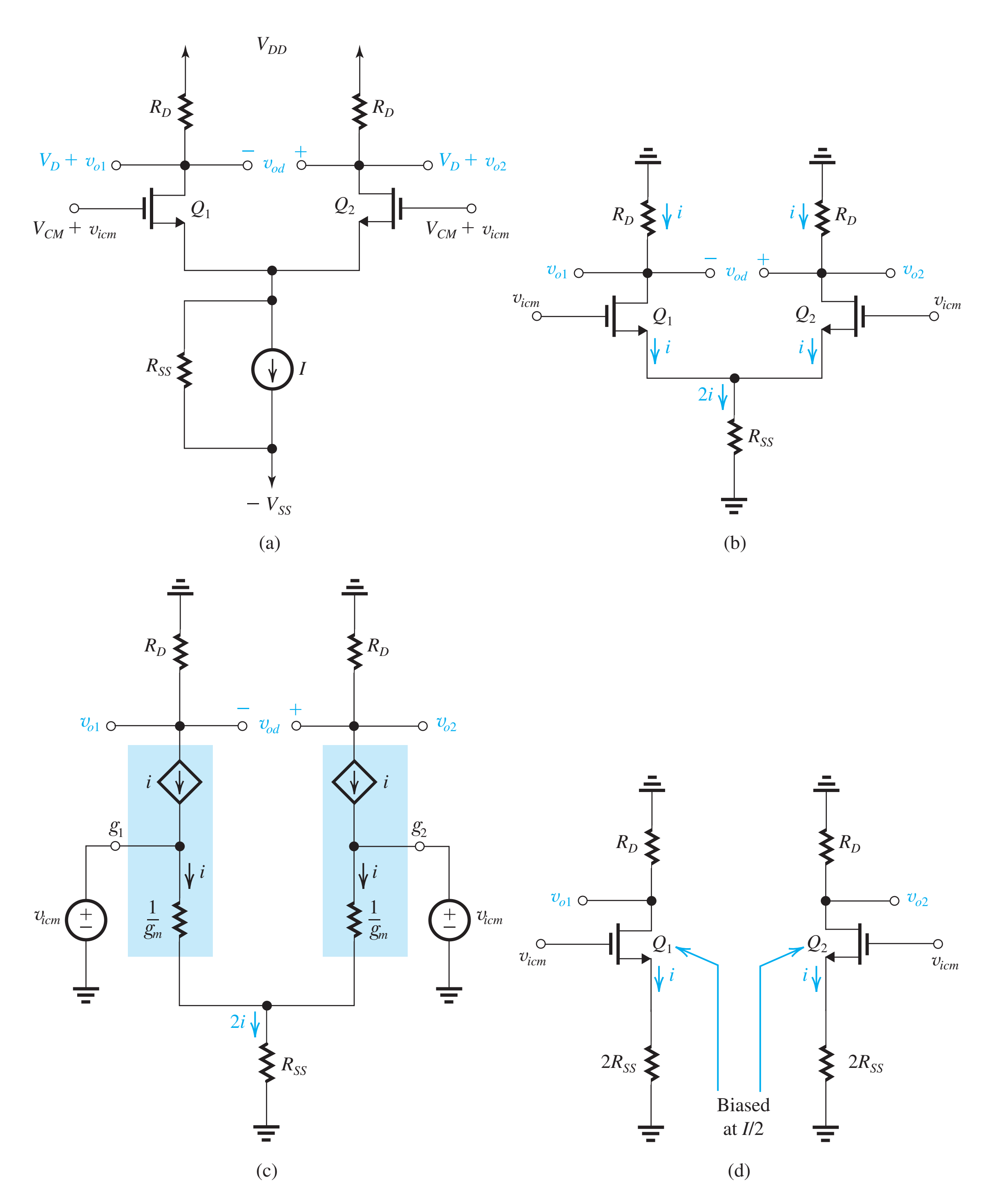


Figure 9.25 (**a**) A MOS differential amplifier with a common-mode input signal v_{icm} superimposed on the input dc common-mode voltage V_{CM}. (**b**) The amplifier circuit prepared for small-signal analysis. (**c**) The amplifier circuit with the transistors replaced with their T model and r_o neglected. (**d**) The circuit in (**b**) split into its two halves; each half is said to be a "CM half-circuit."

Thus the circuit still rejects common-mode signals! Unfortunately, however, this will not be the case if the circuit is not perfectly symmetrical, as we shall now show.

Before proceeding further, it is useful to observe that all the above results can be obtained by considering only half the differential amplifier. Figure 9.25(d) shows the two half-circuits of the differential amplifier that apply for common-mode analysis. To see the equivalence, observe that each of the two half-circuits indeed carries a current i given by Eq. (9.72) and the voltages at the source terminals are equal ($v_s = 2iR_{SS}$). Thus the two sources can be joined, returning the circuit to the original form in Fig. 9.25(b). Each of the circuits in Fig. 9.25(d) is known as the **common-mode half-circuit**. Note the difference between the CM half-circuit and the differential half-circuit.

Effect of R_D Mismatch When the two drain resistances exhibit a mismatch ΔR_D, as they inevitably do, the common-mode voltages at the two drains will no longer be equal. Rather, if the load of Q_1 is R_D and that of Q_2 is $(R_D + \Delta R_D)$ the drain signal voltages arising from v_{icm} will be

$$v_{o1} \simeq -\frac{R_D}{2R_{SS}} v_{icm} \tag{9.76}$$

and

$$v_{o2} \simeq -\frac{R_D + \Delta R_D}{2R_{SS}} v_{icm} \tag{9.77}$$

Thus,

$$v_{od} = v_{o2} - v_{o1} = -\frac{\Delta R_D}{2R_{SS}} v_{icm} \tag{9.78}$$

and we can find the **common-mode gain** A_{cm} as

$$A_{cm} \equiv \frac{v_{od}}{v_{icm}} = -\frac{\Delta R_D}{2R_{SS}} \tag{9.79}$$

which can be expressed in the alternate form

$$A_{cm} = -\left(\frac{R_D}{2R_{SS}}\right)\left(\frac{\Delta R_D}{R_D}\right) \tag{9.80}$$

It follows that a mismatch in the drain resistances causes the differential amplifier to have a finite common-mode gain. Thus, a portion of the interference or noise signal v_{icm} will appear as a component of v_{od}. A measure of the effectiveness of the differential amplifier in amplifying differential-mode signals and rejecting common-mode interference is the ratio of the magnitude of its differential gain $|A_d|$ to the magnitude of its common-mode gain $|A_{cm}|$. This ratio is termed **common-mode rejection ratio (CMRR)**. Thus,

$$\text{CMRR} \equiv \frac{|A_d|}{|A_{cm}|} \tag{9.81}$$

and is usually expressed in decibels,

$$\text{CMRR(dB)} = 20 \log \frac{|A_d|}{|A_{cm}|} \tag{9.82}$$

For the case of a MOS differential amplifier with drain resistances R_D that exhibit a mismatch ΔR_D, the CMRR can be found as the ratio of A_d in Eq. (9.35) to A_{cm} in Eq. (9.79), thus

$$\text{CMRR} = (2g_m R_{SS}) \Big/ \left(\frac{\Delta R_D}{R_D} \right) \tag{9.83}$$

It follows that to obtain a high CMRR, we should utilize a bias current source with a high output resistance R_{SS}, and we should strive to obtain a high degree of matching between the drain resistances (i.e., keep $\Delta R_D/R_D$ small).

EXERCISE

9.11 A MOS differential pair operated at a bias current of 0.8 mA employs transistors with $W/L = 100$ and $\mu_n C_{ox} = 0.2\ \text{mA/V}^2$, using $R_D = 5\ \text{k}\Omega$ and $R_{SS} = 25\ \text{k}\Omega$. Find the differential gain, the common-mode gain when the drain resistances have a 1% mismatch, and the CMRR.
Ans. 20 V/V; 0.001 V/V; 86 dB

Effect of g_m Mismatch on CMRR Another possible mismatch between the two halves of the MOS differential pair is a mismatch in g_m of the two transistors. For the purpose of finding the effect of a g_m mismatch on CMRR, let

$$g_{m1} = g_m + \frac{1}{2}\Delta g_m \tag{9.84}$$

$$g_{m2} = g_m - \frac{1}{2}\Delta g_m \tag{9.85}$$

That is,

$$g_{m1} - g_{m2} = \Delta g_m \tag{9.86}$$

Since the circuit is no longer symmetrical, we cannot employ the common-mode half-circuit. Rather, we need to return to the original circuit of Fig. 9.25(a) and replace each of Q_1 and Q_2 with its T equivalent-circuit model. We shall skip the analysis and simply present the result,

$$A_{cm} \simeq \left(\frac{R_D}{2R_{SS}} \right) \left(\frac{\Delta g_m}{g_m} \right) \tag{9.87}$$

and the corresponding CMRR will be

$$\text{CMRR} = (2g_m R_{SS}) \Big/ \left(\frac{\Delta g_m}{g_m} \right) \tag{9.88}$$

Note that both expressions have exactly the same form as the corresponding expressions for the case of R_D mismatch.

Thus, as in that case, to keep CMRR high, we have to use a biasing current source with a high output resistance R_{SS} and, of course, strive to maintain a high degree of matching between Q_1 and Q_2.

EXERCISE

9.12 For the MOS amplifier specified in Exercise 9.11, compute the CMRR resulting from a 1% mismatch in g_m.
Ans. 86 dB

Example 9.4

In this example we consider the design of the current source that supplies the bias current of a MOS differential amplifier. Let it be required to achieve a CMRR of 100 dB and assume that the only source of mismatch between Q_1 and Q_2 is a 2% mismatch in their W/L ratios. Let $I = 200\ \mu\text{A}$ and assume that all transistors are to be operated at $V_{OV} = 0.2$ V. For the 0.18-μm CMOS fabrication process available, $V_A' = 5$ V/μm. If a simple current source is utilized for I, what channel length is required? If a cascode current source is utilized, what channel length is needed for the two transistors in the cascode?

Solution

A mismatch in W/L results in a g_m mismatch that can be found from the expression of g_m:

$$g_m = \sqrt{2(\mu_n C_{ox})\left(\frac{W}{L}\right)I_D} \tag{9.89}$$

It can be seen that an error of 2% in W/L will result in an error in g_m of 1%. That is, the 2% mismatch in the W/L ratios of Q_1 and Q_2 will result in a 1% mismatch in their g_m values. The resulting CMRR can be found from Eq. (9.88), repeated here:

$$\text{CMRR} = (2g_m R_{SS})\Big/\left(\frac{\Delta g_m}{g_m}\right)$$

Now, a 100-dB CMRR corresponds to a ratio of 10^5; thus,

$$10^5 = (2g_m R_{SS})/0.01 \tag{9.90}$$

The value of g_m can be found from

$$g_m = \frac{2I_D}{V_{OV}} = \frac{2 \times (I/2)}{V_{OV}}$$

$$= \frac{2 \times 0.1}{0.2} = 1 \text{ mA/V}$$

Substituting in Eq. (9.90) gives

$$R_{SS} = 500 \text{ k}\Omega$$

Now if the current source is implemented with a single transistor, its r_o must be

$$r_o = R_{SS} = 500 \text{ k}\Omega$$

Thus,

$$\frac{V_A}{I} = 500\ \text{k}\Omega$$

Substituting $I = 200\ \mu\text{A}$, we find the required value of V_A as

$$V_A = 100\ \text{V}$$

Since $V_A = V_A'L = 5L$, the required value of L will be

$$L = 20\ \mu\text{m}$$

which is very large!

Using a cascode current source, we have

$$R_{SS} = (g_m r_o) r_o$$

where

$$g_m = \frac{2I}{V_{OV}} = \frac{2 \times 0.2}{0.2} = 2\ \text{mA/V}$$

Thus,

$$500 = 2 \times r_o^2$$
$$r_o = 15.81\ \text{k}\Omega$$

and the required V_A now becomes

$$15.81 = \frac{V_A}{I} = \frac{V_A}{0.2}$$
$$V_A = 3.16\ \ \text{V}$$

which implies a channel length for each of the two transistors in the cascode of

$$L = \frac{3.16}{V_A'} = \frac{3.16}{5} = 0.63\ \mu\text{m}$$

a considerable reduction from the case of a simple current source, and indeed a practical value.

Differential versus Single-Ended Output The above study of common-mode rejection was predicated on the assumption that the output of the differential amplifier is taken differentially, that is, between the drains of Q_1 and Q_2. In some cases one might decide to take the output single-endedly; that is, between one of the drains and ground. If this is done, the CMRR is reduced dramatically. This can be seen from the above analysis, where the common-mode gain in the absence of mismatches is zero if the output is taken differentially and finite (Eq. 9.74) if the output is taken single-endedly. When mismatches are taken into account, the CM gain for the differential-output case departs from zero but remains much lower than the value obtained for single-ended output (Eq. 9.74).

We conclude that to obtain a large CMRR, the output of the differential amplifier must be taken differentially. The subject of converting the output signal from differential to single-ended without loss of CMRR will be studied in Section 9.5.

EXERCISE

9.13 Show that if the output of the MOS differential amplifier is taken single-endedly, the CMRR is given by:

$$\text{CMRR} = g_m R_{SS}$$

9.3.2 The BJT Case

An exactly similar development applies for studying the common-mode rejection of the BJT differential amplifier. Figure 9.26 shows a bipolar differential amplifier with an input common-mode signal v_{icm}. Here R_{EE} is the output resistance of the bias current source I. We wish to find the voltages that result from v_{icm} at the collectors of Q_1 and Q_2, v_{o1} and v_{o2}, and between the two collectors, v_{od}. Toward that end, we make use of the **common-mode half-circuits** shown in Fig. 9.26(b). The signal v_{o1} that appears at the collector of Q_1 in response to v_{icm} will be

$$v_{o1} = -\frac{\alpha R_C}{r_e + 2R_{EE}} v_{icm} \tag{9.91}$$

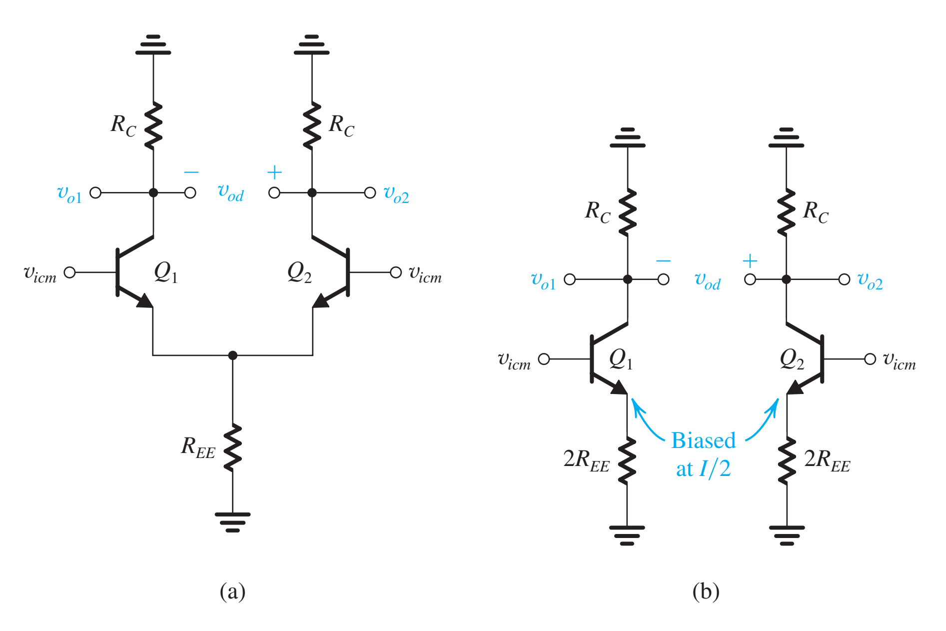


Figure 9.26 **(a)** The differential amplifier fed by a common-mode input signal v_{icm}. **(b)** Equivalent "half-circuits" for common-mode calculations.

Similarly, v_{o2} will be

$$v_{o2} = -\frac{\alpha R_C}{r_e + 2R_{EE}} v_{icm} \tag{9.92}$$

where we have neglected the transistor r_o, for simplicity. The differential output signal v_{od} can be obtained as

$$v_{od} = v_{o2} - v_{o1} = 0$$

Thus, while the voltages at the two collectors will contain common-mode noise or interference components, the output differential voltage will be free from such interference. This condition, however, is based on the assumption of perfect matching between the two sides of the differential amplifier. Any mismatch will result in v_{od} acquiring a component proportional to v_{icm}. For example, a mismatch ΔR_C between the two collector resistances results in

$$A_{cm} \equiv \frac{v_{od}}{v_{icm}} = -\frac{\alpha \Delta R_C}{2R_{EE} + r_e} \tag{9.93}$$

Since $\alpha \simeq 1$, $r_e \ll 2R_{EE}$, Eq. (9.93) can be approximated and written in the form

$$A_{cm} \simeq -\left(\frac{R_C}{2R_{EE}}\right)\left(\frac{\Delta R_C}{R_C}\right) \tag{9.94}$$

The common-mode rejection ratio can now be found from

$$\text{CMRR} = \frac{|A_d|}{|A_{cm}|}$$

together with using Eqs. (9.68) and (9.94), with the result that

$$\text{CMRR} = (2g_m R_{EE}) \Big/ \left(\frac{\Delta R_C}{R_C}\right) \tag{9.95}$$

which is similar in form to the expression for the MOS pair [Eq. (9.83)]. Thus, to obtain a high CMRR, we design the current source to have a large output resistance R_{EE} and strive for close matching of the collector resistances.

Common-Mode Input Resistance The definition of the common-mode input resistance R_{icm} is illustrated in Fig. 9.27(a). Figure 9.27(b) shows the equivalent common-mode half-circuit; its input resistance is $2R_{icm}$. The value of $2R_{icm}$ can be determined by analyzing the circuit of Fig. 9.27(b) while taking r_o into account (because R_{EE} and R_C can be equal to, *or* larger than, r_o). The analysis is straightforward but tedious and can be shown (Problem 9.79) to yield the following result

$$R_{icm} \simeq \beta R_{EE} \frac{1 + R_C/\beta r_o}{1 + \dfrac{R_C + 2R_{EE}}{r_o}} \tag{9.96}$$

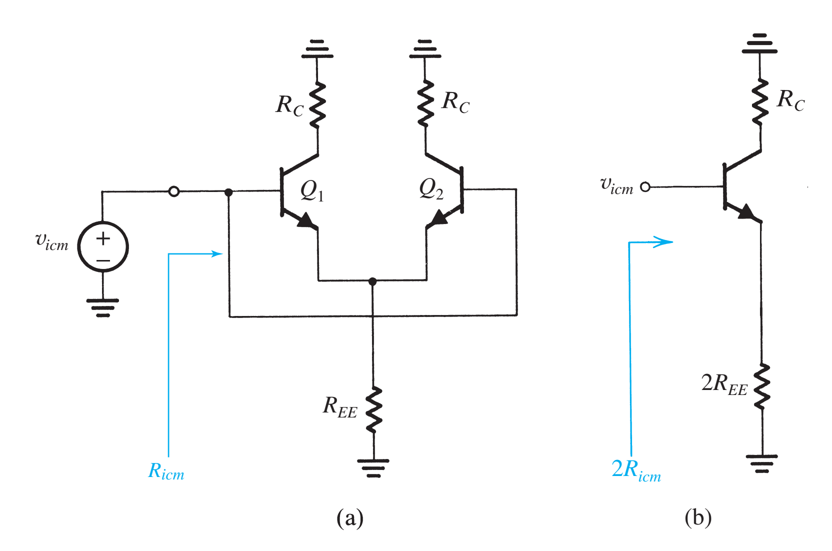


Figure 9.27 **(a)** Definition of the input common-mode resistance R_{icm}. **(b)** The equivalent common-mode half-circuit.

Example 9.5

For the differential amplifier analyzed in Example 9.3, let the bias-current source have an output resistance $R_{EE} = 200\ \text{k}\Omega$. Evaluate:

(a) the worst-case common-mode gain if the two collector resistances are accurate to within $\pm 1\%$.
(b) the CMRR in dB.
(c) the input common-mode resistance (assuming the Early voltage $V_A = 100$ V).

Solution

First we observe that the two emitter resistances R_E will have negligible effect on A_{cm}.

(a) Using Eq. (9.94),

$$|A_{cm}| = \frac{R_C}{2R_{EE}} \frac{\Delta R_C}{R_C}$$

where $\Delta R_C = 0.02 R_C$ in the worst case. Thus,

$$|A_{cm}| = \frac{10}{2 \times 200} \times 0.02 = 5 \times 10^{-4} \text{ V/V}$$

(b)

$$\text{CMRR} = 20 \log \frac{|A_d|}{|A_{cm}|}$$

where from Example 9.3, $|A_d| = 40$, thus

$$\text{CMRR} = 20 \log \frac{40}{5 \times 10^{-4}} = 98 \text{ dB}$$

(c)

$$r_o = \frac{V_A}{I/2} = \frac{100}{0.5} = 200\ \text{k}\Omega$$

Using Eq. (9.96),

$$R_{\text{in}} = 6.6\ \text{M}\Omega$$

EXERCISE

9.14 A bipolar differential amplifier utilizes a simple (i.e., a single CE transistor) current source to supply a bias current I of 200 μA, and simple current-source loads formed by *pnp* transistors. For all transistors, $\beta = 100$ and $|V_A| = 10$ V. Find g_m, R_C, $|A_d|$, R_{id}, R_{EE}, CMRR (if the two load transistors exhibit a 1% mismatch in their r_o's), and R_{icm}. *Hint*: Remember to take into account r_{o1} and r_{o2}.
Ans. 4 mA/V; 100 kΩ; 200 V/V; 50 kΩ, 50 kΩ; 86 dB; 1.67 MΩ

9.4 DC Offset

Because differential amplifiers are directly coupled and have finite gain at dc, they suffer from a number of dc problems. In this section we study some of these.

9.4.1 Input Offset Voltage of the MOS Differential Amplifier

Consider the basic MOS differential amplifier with both inputs grounded, as shown in Fig. 9.28(a). If the two sides of the differential pair were perfectly matched (i.e., Q_1 and Q_2 identical and $R_{D1} = R_{D2} = R_D$), then current I would split equally between Q_1 and Q_2, and V_O would be zero. But practical circuits exhibit mismatches that result in a dc output voltage V_O even with both inputs grounded. We call V_O the **output dc offset voltage**. More commonly, we divide V_O by the differential gain of the amplifier, A_d, to obtain a quantity known as the **input offset voltage**, V_{OS},

$$V_{OS} = V_O/A_d \tag{9.97}$$



We can see that if we apply a voltage $-V_{OS}$ between the input terminals of the differential amplifier, then the output voltage will be reduced to zero [see Fig. 9.28(b)]. This observation gives rise to the usual definition of the input offset voltage. It should be noted, however, that since the offset voltage is a result of device mismatches, its polarity is not known a priori.

Three factors contribute to the dc offset voltage of the MOS differential pair: mismatch in load resistances, mismatch in W/L, and mismatch in V_t. We shall consider the three contributing factors one at a time.

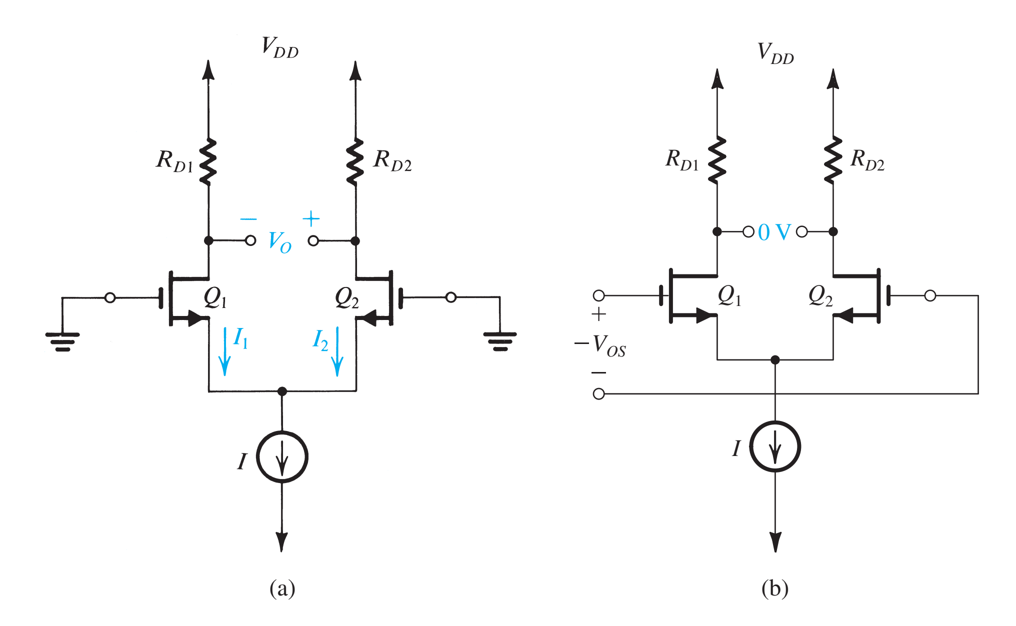


Figure 9.28 **(a)** The MOS differential pair with both inputs grounded. Owing to device and resistor mismatches, a finite dc output voltage V_O results. **(b)** Application of a voltage equal to the input offset voltage V_{OS} to the input terminals with opposite polarity reduces V_O to zero.

For the differential pair shown in Fig. 9.28(a) consider first the case where Q_1 and Q_2 are perfectly matched but R_{D1} and R_{D2} show a mismatch ΔR_D; that is,

$$R_{D1} = R_D + \frac{\Delta R_D}{2} \tag{9.98}$$

$$R_{D2} = R_D - \frac{\Delta R_D}{2} \tag{9.99}$$

Because Q_1 and Q_2 are matched, the current I will split equally between them. Nevertheless, because of the mismatch in load resistances, the output voltages V_{D1} and V_{D2} will be

$$V_{D1} = V_{DD} - \frac{I}{2}\left(R_D + \frac{\Delta R_D}{2}\right)$$

$$V_{D2} = V_{DD} - \frac{I}{2}\left(R_D - \frac{\Delta R_D}{2}\right)$$

Thus the differential output voltage V_O will be

$$\begin{aligned} V_O &= V_{D2} - V_{D1} \\ &= \left(\frac{I}{2}\right)\Delta R_D \end{aligned} \tag{9.100}$$

The corresponding input offset voltage is obtained by dividing V_O by the gain $g_m R_D$ and substituting for g_m from Eq. (9.30). The result is

$$V_{OS} = \left(\frac{V_{OV}}{2}\right)\left(\frac{\Delta R_D}{R_D}\right) \tag{9.101}$$

Thus the offset voltage is directly proportional to V_{OV} and, of course, to $\Delta R_D/R_D$. As an example, consider a differential pair in which the two transistors are operating at an overdrive voltage of 0.2 V and each drain resistance is accurate to within ±1%. It follows that the worst-case resistor mismatch will be

$$\frac{\Delta R_D}{R_D} = 0.02$$

and the resulting input offset voltage will be

$$|V_{OS}| = 0.1 \times 0.02 = 2 \text{ mV}$$

Next, consider the effect of a mismatch in the W/L ratios of Q_1 and Q_2, expressed as

$$\left(\frac{W}{L}\right)_1 = \frac{W}{L} + \frac{1}{2}\Delta\left(\frac{W}{L}\right) \quad (9.102)$$

$$\left(\frac{W}{L}\right)_2 = \frac{W}{L} - \frac{1}{2}\Delta\left(\frac{W}{L}\right) \quad (9.103)$$

Such a mismatch causes the current I to no longer divide equally between Q_1 and Q_2. Rather, because $V_{GS1} = V_{GS2}$, the current conducted by each of Q_1 and Q_2 will be proportional to its W/L ratio, and we can easily show that

$$I_1 = \frac{I}{2}\left[1 + \frac{\Delta(W/L)}{2(W/L)}\right] \quad (9.104)$$

$$I_2 = \frac{I}{2}\left[1 - \frac{\Delta(W/L)}{2(W/L)}\right] \quad (9.105)$$

Dividing the current difference,

$$\frac{I}{2}\frac{\Delta(W/L)}{(W/L)}$$

by g_m gives the input offset voltage (due to the mismatch in W/L values).[2] Thus

$$V_{OS} = \left(\frac{V_{OV}}{2}\right)\left(\frac{\Delta(W/L)}{(W/L)}\right) \quad (9.106)$$

Here again we note that V_{OS}, resulting from a (W/L) mismatch, is proportional to V_{OV} and, as expected, $\Delta(W/L)$.

Finally, we consider the effect of a mismatch ΔV_t between the two threshold voltages,

$$V_{t1} = V_t + \frac{\Delta V_t}{2} \quad (9.107)$$

$$V_{t2} = V_t - \frac{\Delta V_t}{2} \quad (9.108)$$

The current I_1 will be given by

$$I_1 = \frac{1}{2}k_n'\frac{W}{L}\left(V_{GS} - V_t - \frac{\Delta V_t}{2}\right)^2$$

$$= \frac{1}{2}k_n'\frac{W}{L}(V_{GS} - V_t)^2\left(1 - \frac{\Delta V_t}{2(V_{GS} - V_t)}\right)^2$$

[2] We are skipping a step in the derivation: Rather than multiplying the current difference by R_C and dividing the resulting output offset by $A_d = g_m R_C$, we are simply dividing the current difference by g_m.

which, for $\Delta V_t \ll 2(V_{GS} - V_t)$ (that is, $\Delta V_t \ll 2\ V_{OV}$), can be approximated as

$$I_1 \simeq \frac{1}{2}k_n' \frac{W}{L}(V_{GS} - V_t)^2\left(1 - \frac{\Delta V_t}{V_{GS} - V_t}\right)$$

Similarly,

$$I_2 \simeq \frac{1}{2}k_n' \frac{W}{L}(V_{GS} - V_t)^2\left(1 + \frac{\Delta V_t}{V_{GS} - V_t}\right)$$

We recognize that

$$\frac{1}{2}k_n' \frac{W}{L}(V_{GS} - V_t)^2 = \frac{I}{2}$$

and the current increment (decrement) in Q_2 (Q_1) is

$$\Delta I = \frac{I}{2}\frac{\Delta V_t}{V_{GS} - V_t} = \frac{I}{2}\frac{\Delta V_t}{V_{OV}}$$

Dividing the current difference $2\Delta I$ by g_m gives the input offset voltage (due to ΔV_t). Thus,

$$V_{OS} = \Delta V_t \qquad (9.109)$$

a very logical result! For modern MOS technology ΔV_t can be as high as a few mV. Finally, we note that since the three sources for offset voltage are not correlated, an estimate of the total input offset voltage can be found as

$$V_{OS} = \sqrt{\left(\frac{V_{OV}}{2}\frac{\Delta R_D}{R_D}\right)^2 + \left(\frac{V_{OV}}{2}\frac{\Delta(W/L)}{W/L}\right)^2 + (\Delta V_t)^2} \qquad (9.110)$$

EXERCISE

9.15 For the MOS differential pair specified in Exercise 9.4, find the three components of the input offset voltage. Let $\Delta R_D/R_D = 2\%$, $\Delta(W/L)/(W/L) = 2\%$, and $\Delta V_t = 2$ mV. Use Eq. (9.110) to obtain an estimate of the total V_{OS}.
Ans. 2 mV; 2 mV; 2 mV; 3.5 mV

9.4.2 Input Offset Voltage of the Bipolar Differential Amplifier

The offset voltage of the bipolar differential pair shown in Fig. 9.29(a) can be determined in a manner analogous to that used above for the MOS pair. Note, however, that in the bipolar case there is no analog to the V_t mismatch of the MOSFET pair. Here the output offset results from mismatches in the load resistances R_{C1} and R_{C2} and from junction area, β, and other mismatches in Q_1 and Q_2. Consider first the effect of the load mismatch. Let

$$R_{C1} = R_C + \frac{\Delta R_C}{2} \qquad (9.111)$$

$$R_{C2} = R_C - \frac{\Delta R_C}{2} \qquad (9.112)$$

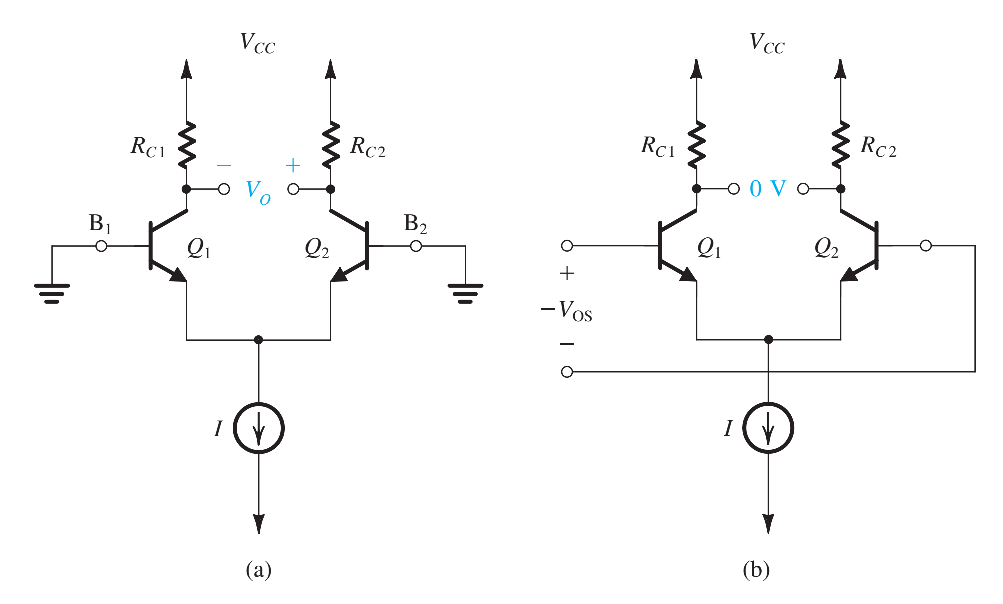


Figure 9.29 **(a)** The BJT differential pair with both inputs grounded. Device mismatches result in a finite dc output V_O. **(b)** Application of the input offset voltage $V_{OS} \equiv V_O/A_d$ to the input terminals with opposite polarity reduces V_O to zero.

and assume that Q_1 and Q_2 are perfectly matched. It follows that current I will divide equally between Q_1 and Q_2, and thus

$$V_{C1} = V_{CC} - \left(\frac{\alpha I}{2}\right)\left(R_C + \frac{\Delta R_C}{2}\right)$$

$$V_{C2} = V_{CC} - \left(\frac{\alpha I}{2}\right)\left(R_C - \frac{\Delta R_C}{2}\right)$$

Thus the output voltage will be

$$V_O = V_{C2} - V_{C1} = \alpha\left(\frac{I}{2}\right)(\Delta R_C)$$

and the input offset voltage will be

$$V_{OS} = \frac{\alpha(I/2)(\Delta R_C)}{A_d} \tag{9.113}$$

Substituting $A_d = g_m R_C$ and

$$g_m = \frac{\alpha I/2}{V_T}$$

gives

$$|V_{OS}| = V_T\left(\frac{\Delta R_C}{R_C}\right) \tag{9.114}$$

An important point to note is that in comparison to the corresponding expression for the MOS pair (Eq. 9.101) here the offset is proportional to V_T rather than $V_{OV}/2$. V_T at 25 mV is 3 to 6 times lower than $V_{OV}/2$. Hence bipolar differential pairs exhibit lower offsets than their MOS

counterparts. As an example, consider the situation of collector resistors that are accurate to within ±1%. Then the worst-case mismatch will be

$$\frac{\Delta R_C}{R_C} = 0.02$$

and the resulting input offset voltage will be

$$|V_{OS}| = 25 \times 0.02 = 0.5 \text{ mV}$$

Next consider the effect of mismatches in transistors Q_1 and Q_2. In particular, let the transistors have a mismatch in their emitter–base junction areas. Such an area mismatch gives rise to a proportional mismatch in the scale currents I_S,

$$I_{S1} = I_S + \frac{\Delta I_S}{2} \tag{9.115}$$

$$I_{S2} = I_S - \frac{\Delta I_S}{2} \tag{9.116}$$

Refer to Fig. 9.29(a) and note that $V_{BE1} = V_{BE2}$. Thus, the current I will split between Q_1 and Q_2 in proportion to their I_S values, resulting in

$$I_{E1} = \frac{I}{2}\left(1 + \frac{\Delta I_S}{2I_S}\right) \tag{9.117}$$

$$I_{E2} = \frac{I}{2}\left(1 - \frac{\Delta I_S}{2I_S}\right) \tag{9.118}$$

It follows that the output offset voltage will be

$$V_O = \alpha\left(\frac{I}{2}\right)\left(\frac{\Delta I_S}{I_S}\right)R_C$$

and the corresponding input offset voltage will be

$$|V_{OS}| = V_T\left(\frac{\Delta I_S}{I_S}\right) \tag{9.119}$$

As an example, an area mismatch of 4% gives rise to $\Delta I_S/I_S = 0.04$ and an input offset voltage of 1 mV. Here again we note that the offset voltage is proportional to V_T rather than to the much larger V_{OV}, which determines the offset of the MOS pair due to $\Delta(W/L)$ mismatch.

Since the two contributions to the input offset voltage are usually not correlated, an estimate of the total input offset voltage can be found as

$$V_{OS} = \sqrt{\left(V_T\frac{\Delta R_C}{R_C}\right)^2 + \left(V_T\frac{\Delta I_S}{I_S}\right)^2}$$

$$= V_T\sqrt{\left(\frac{\Delta R_C}{R_C}\right)^2 + \left(\frac{\Delta I_S}{I_S}\right)^2} \tag{9.120}$$

There are other possible sources for input offset voltage such as mismatches in the values of β and r_o. Some of these are investigated in the end-of-chapter problems. Finally, it should be noted that there is a popular scheme for compensating for the offset voltage. It involves introducing a deliberate mismatch in the values of the two collector resistances such that the differential output voltage is reduced to zero when both input terminals are grounded. Such an **offset-nulling** scheme is explored in Problem 9.81.

9.4.3 Input Bias and Offset Currents of the Bipolar Differential Amplifier

In a perfectly symmetric bipolar differential pair, the two input terminals carry equal dc currents; that is,

$$I_{B1} = I_{B2} = \frac{I/2}{\beta + 1} \tag{9.121}$$

This is the **input bias current** of the differential amplifier.

Mismatches in the amplifier circuit and most importantly a mismatch in β make the two input dc currents unequal. The resulting difference is the **input offset current**, I_{OS}, given as

$$I_{OS} = |I_{B1} - I_{B2}| \tag{9.122}$$

Let

$$\beta_1 = \beta + \frac{\Delta\beta}{2}$$

$$\beta_2 = \beta - \frac{\Delta\beta}{2}$$

then

$$I_{B1} = \frac{I}{2}\frac{1}{\beta + 1 + \Delta\beta/2} \simeq \frac{I}{2}\frac{1}{\beta + 1}\left(1 - \frac{\Delta\beta}{2\beta}\right) \tag{9.123}$$

$$I_{B2} = \frac{I}{2}\frac{1}{\beta + 1 - \Delta\beta/2} \simeq \frac{I}{2}\frac{1}{\beta + 1}\left(1 + \frac{\Delta\beta}{2\beta}\right) \tag{9.124}$$

$$I_{OS} = \frac{I}{2(\beta + 1)}\left(\frac{\Delta\beta}{\beta}\right) \tag{9.125}$$

Formally, the input bias current I_B is defined as follows:

$$I_B \equiv \frac{I_{B1} + I_{B2}}{2} = \frac{I}{2(\beta + 1)} \tag{9.126}$$

Thus

$$I_{OS} = I_B\left(\frac{\Delta\beta}{\beta}\right) \tag{9.127}$$

As an example, a 10% β mismatch results in an offset current that is one-tenth the value of the input bias current.

Finally note that a great advantage of the MOS differential pair is that it does not suffer from a finite input bias current or from mismatches thereof!

EXERCISE

9.16 For a BJT differential amplifier utilizing transistors having $\beta = 100$, matched to 10% or better, and areas that are matched to 10% or better, along with collector resistors that are matched to 2% or better, find V_{OS}, I_B, and I_{OS}. The dc bias current I is 100 μA.

Ans. 2.55 mV; 0.5 μA; 50 nA

9.4.4 A Concluding Remark

We conclude this section by noting that the definitions presented here are identical to those presented in Chapter 2 for op amps. In fact, as will be seen in Chapter 13, it is the input differential stage in an op-amp circuit that primarily determines the op-amp dc offset voltage, input bias and offset currents, and input common-mode range.

9.5 The Differential Amplifier with a Current-Mirror Load

The differential amplifiers we have studied thus far have been of the differential output variety; that is, the output is taken between the two drains (or two collectors) rather than between one of the drains (collectors) and ground. Taking the output differentially has three major advantages:

1. It decreases the common-mode gain and thus increases the common-mode rejection ratio (CMRR). Recall that while the drain (collector) voltages change somewhat in response to a common-mode input signal, the difference between the drain (collector) voltages remains essentially zero except for a small change due to the mismatches inevitably present in the circuit.
2. It decreases the input offset voltage.
3. It increases the differential gain by a factor of 2 (6 dB) because the output is the difference between two voltages of equal magnitude and opposite sign.

These advantages are sufficiently compelling that at least the first stage in an IC amplifier such as an op amp is **differential-in**, **differential-out**. The differential transmission of the signal on the chip also minimizes its susceptibility to corruption with noise and interference, which usually occur in a common-mode fashion. Nevertheless, it is usually required at some point to convert the signal from differential to single-ended; for instance, to connect it to an off-chip load. Figure 9.30 shows a block diagram of a three-stage amplifier in which the first two stages are of the differential-in, differential-out type, and the third has a single-ended output, that is, an output that is referenced to ground. We now address the question of conversion from differential to single-ended.

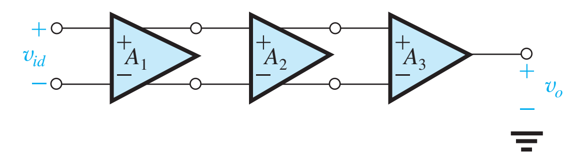


Figure 9.30 A three-stage amplifier consisting of two differential-in, differential-out stages, A_1 and A_2, and a differential-in, single-ended-out stage A_3.

9.5.1 Differential-to-Single-Ended Conversion

Figure 9.31 illustrates the simplest, most basic approach for differential-to-single-ended conversion. It consists of simply ignoring the drain current signal of Q_1 and eliminating its drain resistor altogether, and taking the output between the drain of Q_2 and ground. The

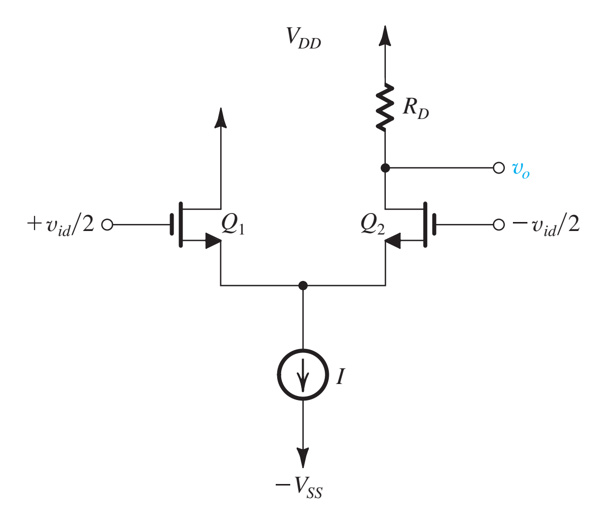


Figure 9.31 A simple but inefficient approach for differential-to-single-ended conversion.

obvious drawback of this scheme is that we lose a factor of 2 (or 6 dB) in gain as a result of "wasting" the drain signal current of Q_1. A much better approach would be to find a way of utilizing the drain current signal of Q_1, and that is exactly what the circuit we are about to discuss accomplishes.

9.5.2 The Current-Mirror-Loaded MOS Differential Pair

Figure 9.32(a) shows a MOS differential pair formed by transistors Q_1 and Q_2, loaded by a current mirror formed by transistors Q_3 and Q_4. To see how this circuit operates, consider first the quiescent or equilibrium state with the two input terminals connected to a dc voltage equal to the common-mode equilibrium value, in this case 0 V, as shown in Fig. 9.32(b). Assuming perfect matching, the bias current I divides equally between Q_1 and Q_2. The drain current of Q_1, $I/2$, is fed to the input transistor of the mirror, Q_3. Thus, a replica of this current is provided by the output transistor of the mirror, Q_4. Observe that at the output node the two currents $I/2$ balance each other out, leaving a zero current to flow out to the next stage or to a load (not shown). This is obviously the desired result! Further, if Q_4 is perfectly matched to Q_3, its drain voltage will track the voltage at the drain of Q_3; thus in equilibrium the voltage at the output will be $V_{DD} - V_{SG3}$. It should be noted, however, that in practical implementations, there will always be mismatches, resulting in a net dc current at the output. In the absence of a load resistance, this current will flow into the output resistances of Q_2 and Q_4 and thus can cause a large deviation in the output voltage from the ideal value. Therefore, this circuit is always designed so that the dc bias voltage at the output node is defined by a feedback circuit rather than by simply relying on the matching of Q_4 and Q_3. We shall see how this is done later.

Next, consider the circuit with a differential input signal v_{id} applied to the input, as shown in Fig. 9.32(c). Since we are now investigating the small-signal operation of the circuit, we have removed the dc supplies (including the current source I). Also, for the time being let us ignore r_o of all transistors. As Fig. 9.32(c) shows, a current i flows through Q_1 and Q_2, given by $i = v_{id}\Big/\left(\dfrac{2}{g_m}\right)$. Thus, transistor Q_1 will conduct a drain signal current $i = g_{m1}v_{id}/2$, and transistor Q_2 will conduct an equal but opposite current i. The drain signal current i of Q_1 is

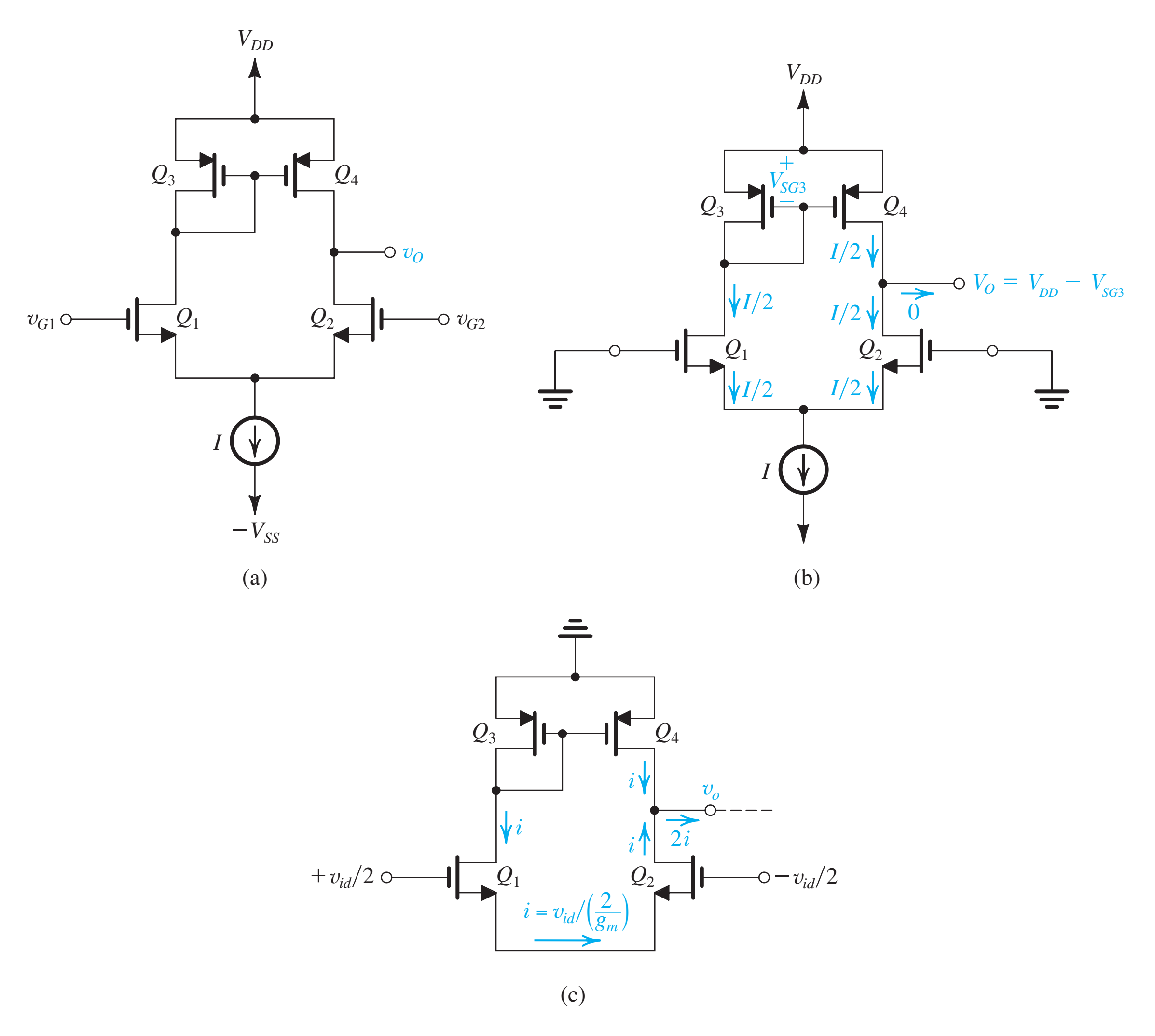


Figure 9.32 **(a)** The current-mirror-loaded MOS differential pair. **(b)** The circuit at equilibrium assuming perfect matching. **(c)** The circuit with a differential input signal applied and neglecting the r_o of all transistors.

fed to the input of the Q_3–Q_4 mirror, which responds by providing a replica in the drain of Q_4. Now, at the output node we have two currents, each equal to i, which sum together to provide an output current $2i$. It is this factor of 2, which is a result of the current-mirror action, that makes it possible to convert the signal to single-ended form (i.e., between the output node and ground) with no loss of gain! If a load resistance is connected to the output node, the current $2i$ flows through it and thus determines the output voltage v_o. In the absence of a load resistance, the output voltage is determined by the output current $2i$ and the output resistance of the circuit, as we shall shortly see.

Before immersing ourselves in detailed analysis of the circuit, it is important to understand the essence of its operation: For dc quantities and common-mode inputs, the current-mirror load produces an output current in the drain of Q_4 that *cancels* the current of Q_2. On the other hand, for differential input signals, the output current of the mirror *adds* to the current of Q_2.

9.5.3 Differential Gain of the Current-Mirror-Loaded MOS Pair

As we learned in Chapter 8, the output resistance r_o of the transistor plays a significant role in the operation of active-loaded amplifiers. Therefore, we shall now take r_o into account and derive an expression for the differential gain v_o/v_{id} of the current-mirror-loaded MOS differential pair. Toward that end, we first observe that the circuit is *not* symmetrical: While the drain of Q_1 sees the small resistance of the diode-connected transistor Q_3 (approximately equal to $1/g_{m3}$), the drain of Q_2 sees the much larger output resistance of Q_4 (r_{o4}). Thus, a virtual ground will *not* develop at the common sources[3] and we cannot use the differential half-circuit technique.

Our approach will be to represent the output of the circuit in Fig. 9.32(c) by the general equivalent circuit shown in Fig. 9.33. Here G_m is the short-circuit transconductance and R_o is the output resistance. In the following, we will show that

$$G_m = g_{m1,2} \tag{9.128}$$

where $g_{m1,2}$ is the transconductance of each of Q_1 and Q_2. We will also show that

$$R_o = r_{o2} \parallel r_{o4} \tag{9.129}$$

In other words, we shall have two intuitively appealing results: The short-circuit transconductance of the circuit is equal to g_m of each of the two transistors of the differential pair, and the output resistance is the parallel equivalent of the output resistances of Q_2 and Q_4. Thus, the open-circuit differential voltage gain can be found as

$$A_d \equiv \frac{v_o}{v_{id}} = G_m R_o = g_{m1,2}(r_{o2} \parallel r_{o4}) \tag{9.130}$$

Writing $g_{m1,2}$ simply as g_m, and for the case $r_{o2} = r_{o4} = r_o$,

$$A_d = \frac{1}{2} g_m r_o = \frac{1}{2} A_0 \tag{9.131}$$

where A_0 is the intrinsic gain of the MOS transistor.

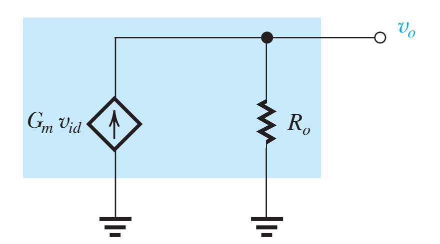


Figure 9.33 Output equivalent circuit of the amplifier in Fig. 9.32(a) for differential input signals.

[3]The qualitative description of circuit operation above implied that a virtual ground develops at the MOSFET sources. That was the case because we were neglecting r_o of all transistors.

EXERCISE

9.17 A current-mirror-loaded MOS differential amplifier of the type shown in Fig. 9.32(a) is specified as follows: $(W/L)_n = 100$, $(W/L)_p = 200$, $\mu_n C_{ox} = 2\mu_p C_{ox} = 0.2\text{ mA/V}^2$, $V_{An} = |V_{Ap}| = 20\text{ V}$, and $I = 0.8$ mA. Calculate G_m, R_o, and A_d.
Ans. 4 mA/V; 25 kΩ; 100 V/V

Derivation of the Short-Circuit Transconductance, G_m Figure 9.34(a) shows the current-mirror-loaded MOS amplifier with the output terminal short-circuited to ground. Our purpose is to determine the short-circuit transconductance

$$G_m \equiv \frac{i_o}{v_{id}}$$

We note that short-circuiting the output terminal makes the circuit nearly balanced. This is because the drain of Q_1 sees the small resistance of the diode-connected transistor Q_3, and now the drain of Q_2 sees a short circuit. It follows that the voltage at the MOSFET sources will be approximately zero. Now, replacing each of the four transistors with its hybrid-π model and noting that for the diode-connected transistor Q_3, the model reduces to a resistance $(1/g_{m3} \| r_{o3})$, we obtain the equivalent circuit shown in Fig. 9.34(b). The short-circuit output current i_o can be found by writing a node equation at the output and noting that the currents in r_{o2} and r_{o4} are zero; thus

$$i_o = g_{m2}\left(\frac{v_{id}}{2}\right) - g_{m4}\, v_{gs4} \tag{9.132}$$

Next, we note that

$$v_{gs4} = v_{gs3} \tag{9.133}$$

and v_{gs3} can be obtained from a node equation at d_1 as

$$v_{gs3} = -g_{m1}\left(\frac{v_{id}}{2}\right)\left(\frac{1}{g_{m3}} \| r_{o3} \| r_{o1}\right)$$

which for the usual case of $\frac{1}{g_{m3}} \ll r_{o3}, r_{o1}$, reduces to

$$v_{gs3} \simeq -\frac{g_{m1}}{g_{m3}}\left(\frac{v_{id}}{2}\right) \tag{9.134}$$

Combining Eqs. (9.132) to (9.134) and substituting $g_{m3} = g_{m4}$ and $g_{m1} = g_{m2} = g_m$ gives

$$i_o = g_m v_{id}$$

from which G_m is found to be

$$G_m = g_m$$

as expected.

Derivation of the Output Resistance R_o Figure 9.35 shows the circuit[4] for determining the output resistance R_o. Observe that we have set v_{id} to zero, resulting in the ground

[4]Note that rather than replacing each transistor with its small-signal model, we are, for simplicity, using the models implicitly. Thus we have "pulled r_o out" of each transistor and shown it separately so that the drain current becomes $g_m v_{gs}$.

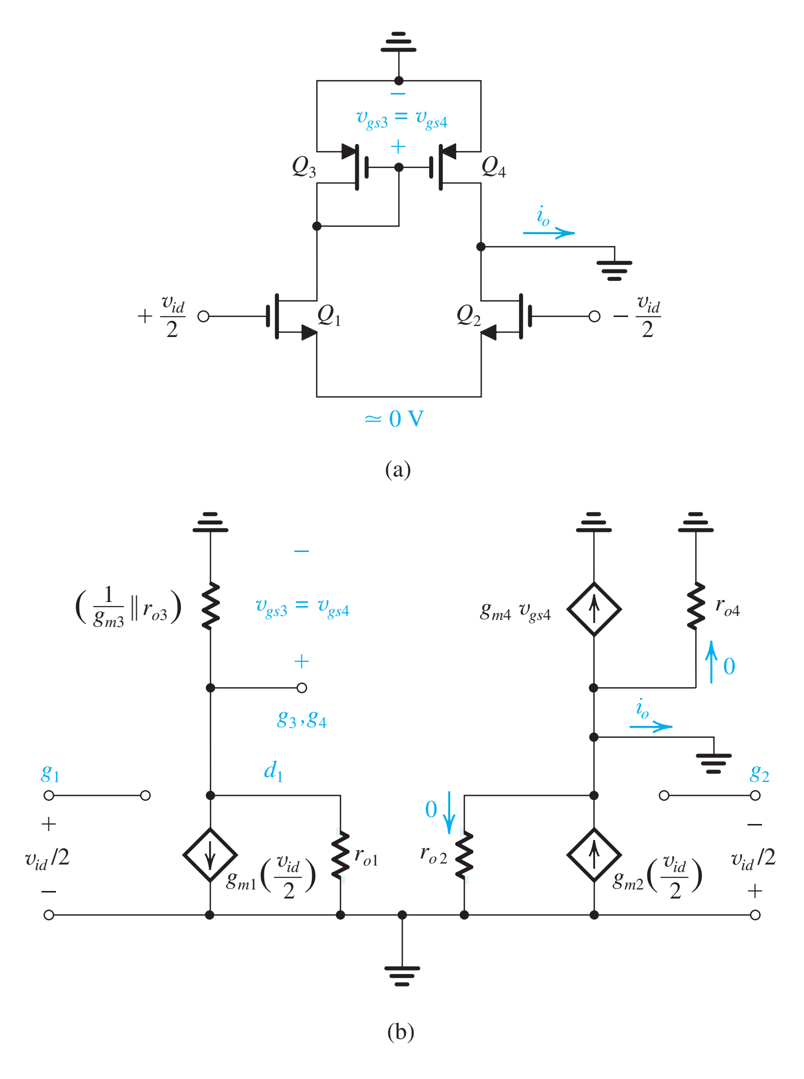


Figure 9.34 Derivation of the short-circuit transconductance $G_m \equiv i_o/v_{id}$.

connections at the gates of Q_1 and Q_2. We have applied a test voltage v_x in order to determine R_o,

$$R_o \equiv \frac{v_x}{i_x}$$

Analysis of this circuit is considerably simplified by observing the current transmission around the circuit by simply following the circled numbers. The current i that enters Q_2 must exit at its source. It then enters Q_1, exiting at the drain to feed the Q_3–Q_4 mirror. Since for the diode-connected transistor Q_3, $1/g_{m3}$ is much smaller than r_{o3}, most of the current i flows into the drain proper of Q_3. The mirror responds by providing an equal current i in the drain of Q_4. The relationship between i and v_x can be determined by observing that at the output node

$$i = v_x/R_{o2}$$

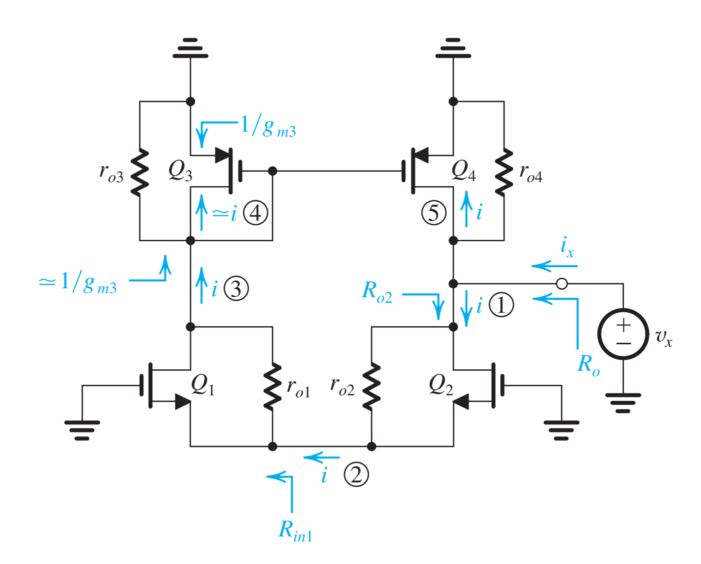


Figure 9.35 Circuit for determining R_o. The circled numbers indicate the order of the analysis steps.

where R_{o2} is the output resistance of Q_2. Now, Q_2 is a CG transistor and has in its source lead the input resistance R_{in1} of the CG transistor Q_1. Noting that the load resistance of Q_1 is $[(1/g_{m3}) \| r_{o3}]$, which is approximately $1/g_{m3}$, we can obtain R_{in1} by using the expression for the input resistance of a CG transistor (Eq. 8.53),

$$R_{\text{in1}} = \frac{r_{o1} + R_L}{g_{m1} r_{o1}}$$
$$= \frac{1}{g_{m1}} + \frac{1/g_{m3}}{g_{m1} r_{o1}} \simeq \frac{1}{g_{m1}}$$

We then use this value of R_{in1} to determine R_{o2} utilizing the expression in Eq. (8.60) as follows:

$$R_{o2} = R_{\text{in1}} + r_{o2} + g_{m2} r_{o2} R_{\text{in1}}$$
$$= \frac{1}{g_{m1}} + r_{o2} + \left(\frac{g_{m2}}{g_{m1}}\right) r_{o2}$$

which, for $g_{m1} = g_{m2} = g_m$ and $g_{m2} r_{o2} \gg 1$, yields

$$R_{o2} \simeq 2 r_{o2} \tag{9.135}$$

Returning to the output node, we write

$$i_x = i + i + \frac{v_x}{r_{o4}}$$
$$= 2i + \frac{v_x}{r_{o4}} = 2\frac{v_x}{R_{o2}} + \frac{v_x}{r_{o4}}$$

Substituting for R_{o2} from Eq. (9.135), we obtain

$$i_x = 2\frac{v_x}{2r_{o2}} + \frac{v_x}{r_{o4}}$$

Thus,

$$R_o \equiv \frac{v_x}{i_x} = r_{o2} \parallel r_{o4} \tag{9.136}$$

which is the result we stated earlier.

9.5.4 The Bipolar Differential Pair with a Current-Mirror Load

The bipolar version of the active-loaded differential pair is shown in Fig. 9.36(a). The circuit structure and operation are very similar to those of its MOS counterpart except that here we have to contend with the effects of finite β and the resulting finite input resistance at the base, r_π. For the time being, however, we shall ignore the effect of finite β on the dc bias of the four transistors and assume that in equilibrium all transistors are operating at a dc current of $I/2$.

Differential Gain To obtain an expression for the differential gain, we use an approach identical to that employed above for the MOS case. That is, we represent the output of the amplifier with the equivalent circuit shown in Fig. 9.36(b) and show that the short-circuit transconductance G_m is given by

$$G_m = g_{m1,2} \tag{9.137}$$

where $g_{m1,2}$ denotes g_m of each of Q_1 and Q_2, and that the output resistance R_o is given by

$$R_o = r_{o2} \parallel r_{o4} \tag{9.138}$$

Both these results are identical to those for the MOS case and can be similarly derived.

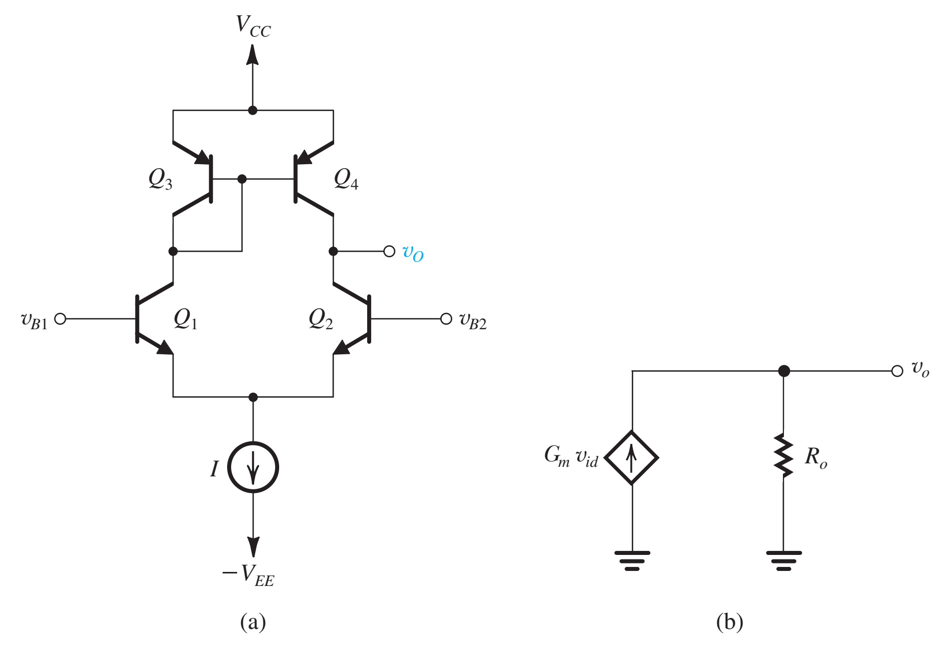


Figure 9.36 **(a)** Current-mirror-loaded bipolar differential pair. **(b)** Small-signal equivalent circuit of the amplifier output when a differential signal $v_{id} \equiv v_{B1} - v_{B2}$ is applied.

Equations (9.137) and (9.138) can now be combined to obtain the differential gain,

$$A_d \equiv \frac{v_o}{v_{id}} = G_m R_o = g_m(r_{o2} \| r_{o4}) \tag{9.139}$$

where $g_m = g_{m1} = g_{m2} \simeq \frac{V_T}{I/2}$, and since $r_{o2} = r_{o4} = r_o$, we can simplify Eq. (9.139) to

$$A_d = \frac{1}{2} g_m r_o \tag{9.140}$$

Although this expression is identical to that found for the MOS circuit, the gain here is much larger because $g_m r_o$ for the BJT is more than an order of magnitude greater than $g_m r_o$ of a MOSFET. The downside, however, lies in the low input resistance of BJT amplifiers. Indeed, from the circuit in Fig. 9.36(a), we can see that the differential input resistance is equal to $2r_\pi$,

$$R_{id} = 2r_\pi \tag{9.141}$$

in sharp contrast to the infinite input resistance of the MOS amplifier. Thus, while the voltage gain realized in a current-mirror-loaded BJT amplifier stage is large, when a subsequent BJT stage is connected to the output, its inevitably low input resistance will drastically reduce the overall voltage gain.

EXERCISE

9.18 For the current-mirror-loaded BJT differential amplifier let $I = 0.8$ mA, $V_A = 100$ V, and $\beta = 160$. Find G_m, R_o, A_d, and R_{id}.

Ans. 16 mA/V; 125 kΩ; 2000 V/V; 20 kΩ

Systematic Input Offset Voltage In addition to the random offset voltages that result from the mismatches inevitably present in the differential amplifier, the current-mirror-loaded bipolar differential pair suffers from a systematic offset voltage that has no counterpart in the MOS version. This is due to the error in the current transfer ratio of the current-mirror load caused by the finite β of the *pnp* transistors that make up the mirror. To see how this comes about, refer to Fig. 9.37. Here the inputs are grounded and the transistors are assumed to be perfectly matched. Thus, the bias current I will divide equally between Q_1 and Q_2, with the result that their two collectors conduct equal currents of $\alpha I/2$. The collector current of Q_1 is fed to the input of the current mirror. From Section 8.2 we know that the current transfer ratio of the mirror is

$$\frac{I_4}{I_3} = \frac{1}{1 + \frac{2}{\beta_P}} \tag{9.142}$$

where β_P is the value of β of the *pnp* transistors Q_3 and Q_4. Thus the collector current of Q_4 will be

$$I_4 = \frac{\alpha I/2}{1 + \frac{2}{\beta_P}} \tag{9.143}$$

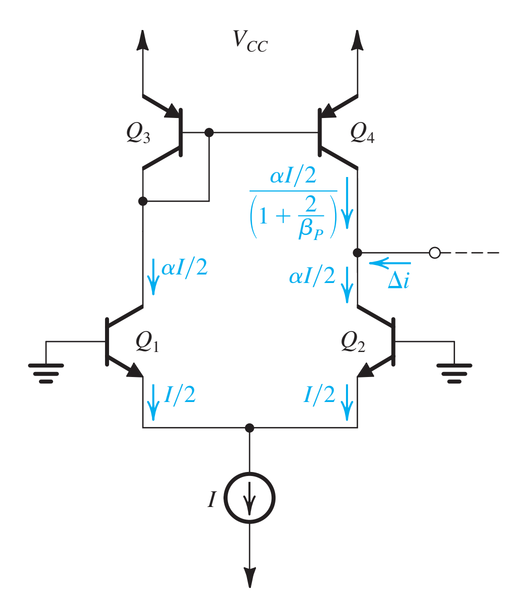


Figure 9.37 The current-mirror-loaded BJT differential pair suffers from a systematic input offset voltage resulting from the error in the current transfer ratio of the current mirror.

which does not exactly balance the collector current of Q_2. It follows that the current difference Δi will flow into the output terminal of the amplifier with

$$\begin{aligned}\Delta i &= \frac{\alpha I}{2} - \frac{\alpha I/2}{1+\dfrac{2}{\beta_P}} \\ &= \frac{\alpha I}{2}\,\frac{2/\beta_P}{1+\dfrac{2}{\beta_P}} \\ &\simeq \frac{\alpha I}{\beta_P}\end{aligned} \tag{9.144}$$

To reduce this output current to zero, an input voltage V_{OS} has to be applied with a value of

$$V_{OS} = -\frac{\Delta i}{G_m}$$

Substituting for Δi from Eq. (9.144) and for $G_m = g_m = (\alpha I/2)/V_T$, we obtain for the input offset voltage the expression

$$V_{OS} = -\frac{\alpha I/\beta_P}{\alpha I/2V_T} = -\frac{2V_T}{\beta_P} \tag{9.145}$$

As an example, for $\beta_P = 50$, $V_{OS} = -1$ mV. To reduce V_{OS}, an improved current mirror such as the Wilson circuit studied in Section 8.6.2 should be used. Such a circuit provides the added advantage of increased output resistance and hence voltage gain. However, to realize the full advantage of the higher output resistance of the active load, the output resistance of the differential pair should be raised by utilizing a cascode stage. Figure 9.38 shows such an arrangement: A folded cascode stage formed by *pnp* transistors Q_3 and Q_4 is utilized to

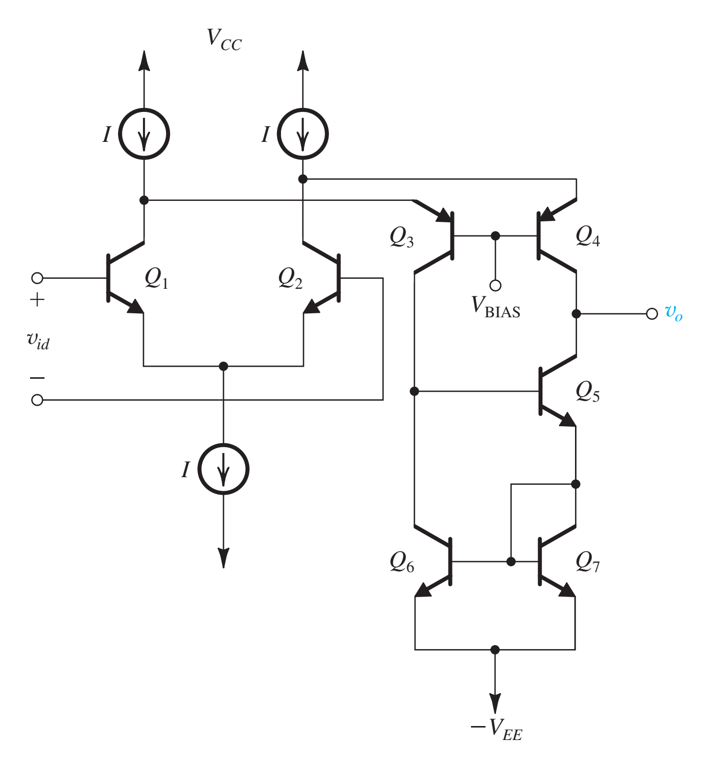


Figure 9.38 An current-mirror-loaded bipolar differential amplifier employing a folded cascode stage (Q_3 and Q_4) and a Wilson current-mirror load (Q_5, Q_6, and Q_7).

raise the output resistance looking into the collector of Q_4 to $\beta_4 r_{o4}$. A Wilson mirror formed by transistors Q_5, Q_6, and Q_7 is used to implement the active load. From Section 8.6.2 we know that the output resistance of the Wilson mirror (i.e., looking into the collector of Q_5) is $\beta_5(r_{o5}/2)$. Thus the output resistance of the amplifier is given by

$$R_o = \left[\beta_4 r_{o4} \parallel \beta_5 \frac{r_{o5}}{2}\right] \tag{9.146}$$

The transconductance G_m remains equal to g_m of Q_1 and Q_2. Thus the differential voltage gain becomes



$$A_d = g_m \left[\beta_4 r_{o4} \parallel \beta_5 \frac{r_{o5}}{2}\right] \tag{9.147}$$

which can be very large. Further examples of improved-performance differential amplifiers will be studied in Chapter 13.

EXERCISE

9.19 Find G_m, R_{o4}, R_{o5}, R_o, and A_d for the differential amplifier in Fig. 9.38 under the following conditions: $I = 1$ mA, $\beta_P = 50$, $\beta_N = 100$, and $V_A = 100$ V.
Ans. 20 mA/V; 10 MΩ; 10 MΩ; 5 MΩ; 10^5 V/V or 100 dB

9.5.5 Common-Mode Gain and CMRR

Although the output is single ended, the current-mirror-loaded differential amplifier has a low common-mode gain (ideally zero) and, correspondingly, a high CMRR (ideally infinite). This is due to the action of the current mirror, whose output current, for common-mode inputs, cancels the current of Q_2 of the differential pair. We have, in fact, seen this in our initial qualitative description of circuit operation.

The current transfer ratio of the mirror, however, will never be exactly unity, and thus the current cancellation at the output node will never be perfect. As a result, the common-mode gain will be finite. We wish to derive an expression for A_{cm}.

Figure 9.39(a) shows the circuit with v_{icm} applied and with the power supplies eliminated except, of course, for the output resistance R_{SS} of the bias-current source I. Although the circuit is not symmetrical and hence we cannot use the common-mode half-circuit, we can split R_{SS} equally between Q_1 and Q_2 as shown in Fig. 9.39(b). It can now be seen that each of Q_1 and Q_2 is a CS transistor with a large source-degeneration resistance $2R_{SS}$.

Each of Q_1 and Q_2 together with their degeneration resistances can be replaced by equivalent circuits composed of a controlled source $G_{mcm}v_{icm}$ and an output resistance $R_{o1,2}$, as shown in Fig. 9.39(c). To determine G_{mcm} we short-circuit the drain to ground, as shown in Fig. 9.39(d) for Q_1. Observe that $2R_{SS}$ and r_{o1} appear in parallel. Thus the voltage at the source terminal can be found from the voltage divider consisting of $1/g_{m1}$ and $(2R_{SS} \| r_{o1})$ as

$$v_s = v_{icm}\frac{(2R_{SS} \| r_{o1})}{(2R_{SS} \| r_{o1}) + (1/g_{m1})}$$

$$\simeq v_{icm}$$

The short-circuit drain current i_o can be seen to be equal to the current through $2R_{SS}$; thus,

$$i_o = \frac{v_s}{2R_{SS}} \simeq \frac{v_{icm}}{2R_{SS}}$$

which leads to

$$G_{mcm} \equiv \frac{i_o}{v_{icm}} = \frac{1}{2R_{SS}} \tag{9.148}$$

The output resistance R_{o1} can be determined using the expression for R_o of a CS transistor with an emitter-degeneration resistance (Eq. 8.60) to obtain

$$R_{o1} = 2R_{SS} + r_{o1} + (g_{m1}r_{o1})(2R_{SS}) \tag{9.149}$$

Similar results can be obtained for Q_2, namely, the same G_{mcm} and an output resistance R_{o2} given by

$$R_{o2} = 2R_{SS} + r_{o2} + (g_{m2}r_{o2})(2R_{SS}) \tag{9.150}$$

Returning to the circuit in Fig. 9.39(c), we see that the current mirror is represented by its input resistance R_{im}, current gain A_m, and output resistance R_{om}. This is a general representation that applies for any current mirror. As current mirrors have relatively low input resistances, R_{im} will be much lower than R_{o1} with the result that the input current of the mirror i_1 will be

$$i_i \simeq G_{mcm}v_{icm} \tag{9.151}$$

The output voltage can be obtained by writing a node equation at the output,

$$v_o = (A_m i_i - G_{mcm}v_{icm})(R_{om} \| R_{o2}) \tag{9.152}$$

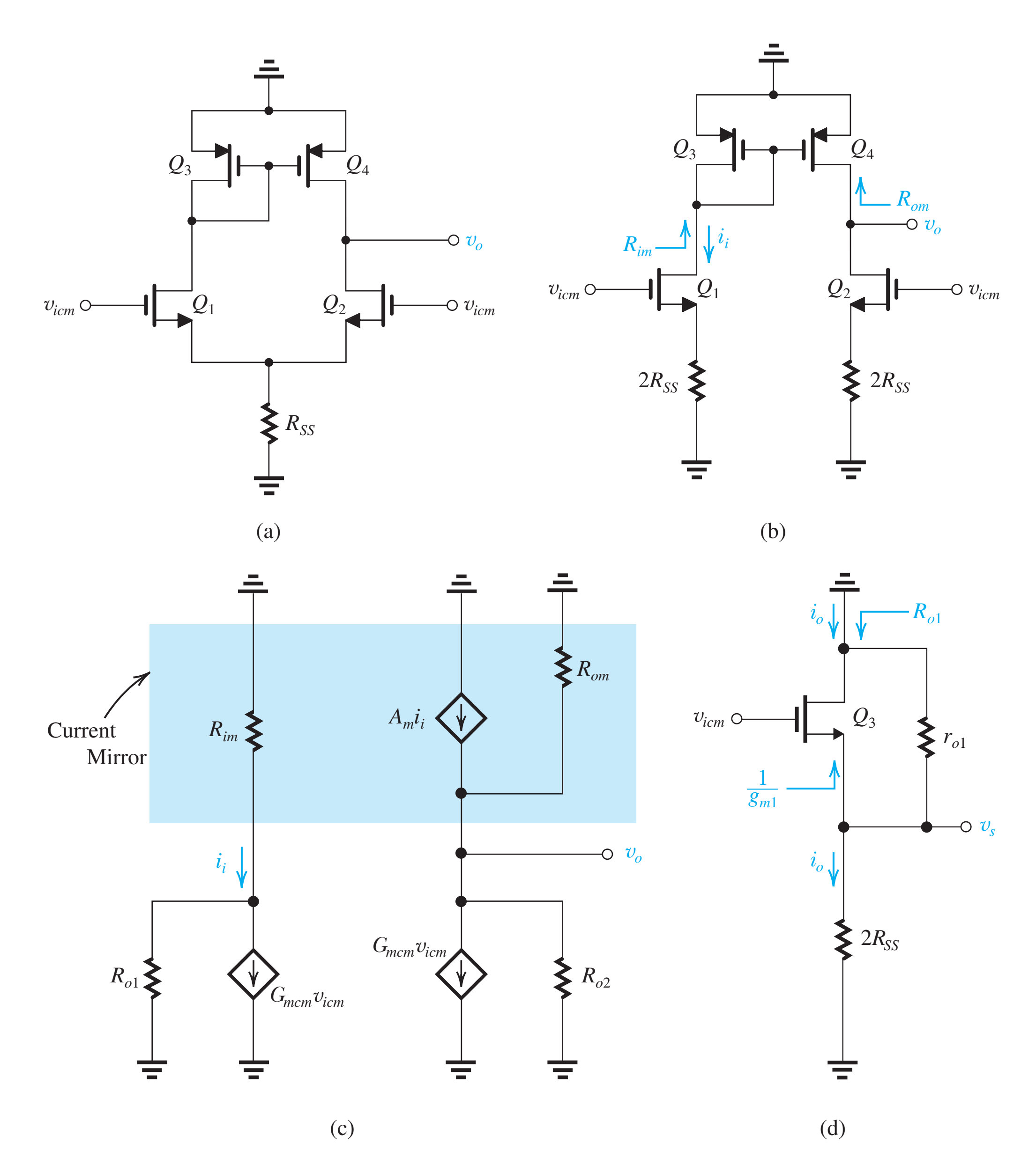


Figure 9.39 Analysis of the current-mirror-loaded MOS differential amplifier to determine its common-mode gain.

Substituting for i_i from Eq. (9.151) results in the following expression for the common-mode gain:

$$A_{cm} \equiv \frac{v_o}{v_{icm}} = -(1 - A_m)G_{mcm}(R_{om} \parallel R_{o2}) \qquad (9.153)$$

This is a general expression that applies for any implementation of the current mirror. The expression makes clear that A_{cm} results from the deviation from unity of the current gain of the mirror.

For the simple current mirror utilized in the circuit of Fig. 9.39(a),

$$R_{im} = \frac{1}{g_{m3}} \parallel r_{o3} \tag{9.154}$$

and

$$R_{om} = r_{o4} \tag{9.155}$$

The current gain A_m can be found as follows

$$A_m i_i = -g_{m4}\, v_{gs4} = -g_{m4}\, v_{gs3}$$

but

$$v_{gs3} = -i_i R_{im}$$

Thus

$$A_m = g_{m4} R_{im}$$

Substituting for R_{im} from (9.154) together with using $g_{m4} = g_{m3}$ results in

$$A_m = 1 \bigg/ \left(1 + \frac{1}{g_{m3}\, r_{o3}}\right) \tag{9.156}$$

Finally, substituting in Eq. (9.153) for A_m from (9.156), for G_{mcm} from (9.148), and for R_{om} from (9.155), and noting that $r_{o4} \ll R_{o2}$, $r_{o4} = r_{o3}$, and $g_m r_{o3} \ll 1$, gives for A_{cm} the expression,

$$A_{cm} \simeq -\frac{1}{2g_{m3}R_{SS}} \tag{9.157}$$

Since R_{SS} is usually large, at least equal to r_o, A_{cm} will be small. The common-mode rejection ratio (CMRR) can now be obtained by utilizing Eqs. (9.130) and (9.157),

$$\text{CMRR} \equiv \frac{|A_d|}{|A_{cm}|} = [g_m(r_{o2} \parallel r_{o4})][2g_{m3}R_{SS}] \tag{9.158}$$

which for $r_{o2} = r_{o4} = r_o$ and $g_{m3} = g_m$ simplifies to

$$\text{CMRR} = (g_m r_o)(g_m R_{SS}) \tag{9.159}$$

We observe that to obtain a large CMRR, we select an implementation of the biasing current source I that features a high output resistance. Such circuits include the cascode current source and the Wilson current source studied in Section 8.6.

EXERCISE

9.20 For the current-mirror-loaded MOS differential amplifier specified in Exercise 9.17, let $R_{SS} = 25\ \text{k}\Omega$. Calculate $|A_{cm}|$ and CMRR. Use the results of Exercise 9.17.
Ans. 0.005 V/V; 20,000 or 86 dB

The Bipolar Case To obtain A_{cm} and CMRR for the BJT circuit of Fig. 9.36(a) we can use the expression in Eq. (9.153) with

$$G_{mcm} = \frac{1}{2R_{EE}} \tag{9.160}$$

$$R_{im} = \frac{1}{g_{m3}} \parallel r_{\pi 3} \parallel r_{o3} \parallel r_{\pi 4} \tag{9.161}$$

which for $r_{o3} \gg r_{\pi 3}$ and $r_{\pi 4} = r_{\pi 3}$ yields

$$R_{im} \simeq \frac{1}{g_{m3}} \parallel \frac{2}{r_{\pi 3}} \tag{9.162}$$

$$R_{om} = r_{o4} \tag{9.163}$$

and

$$A_m = g_{m4} R_{im} \tag{9.164}$$

Assuming $g_{m4} = g_{m3}$ and utilizing $r_{o4} \gg R_{o2}$, we obtain

$$A_{cm} \simeq -\frac{r_{o4}}{2R_{EE}} \frac{\frac{2}{r_{\pi 3}}}{g_{m3} + \frac{2}{r_{\pi 3}}}$$

$$\simeq -\frac{r_{o4}}{2R_{EE}} \frac{2}{\beta_3} = -\frac{r_{o4}}{\beta_3 R_{EE}} \tag{9.165}$$

Using A_d from Eq. (9.139) enables us to obtain the CMRR as

$$\text{CMRR} \equiv \frac{|A_d|}{|A_{cm}|} = g_m (r_{o2} \parallel r_{o4}) \left(\frac{\beta_3 R_{EE}}{r_{o4}} \right) \tag{9.166}$$

For $r_{o2} = r_{o4} = r_o$,

$$\text{CMRR} = \frac{1}{2} \beta_3 g_m R_{EE} \tag{9.167}$$

from which we observe that to obtain a large CMRR, the circuit implementing the bias current source should have a large output resistance R_{EE}. This is possible with, say, a Wilson current mirror (Section 8.6.2).

In conclusion, it is useful to reflect once more on the origin of the finite common-mode gain: It is simply due to the current transmission error introduced by the current-mirror load. In the case of the MOS circuit, this error is due to the finite r_{o3}; in the case of the bipolar mirror, the error is due to the finite β.

EXERCISE

9.21 For the current-mirror-loaded BJT differential amplifier specified in Exercise 9.18, find R_{EE}, A_{cm}, and CMRR. Use the results of Exercise 9.18.
Ans. 125 kΩ; 0.0125 V/V; 160,000, or 104 dB

9.6 Multistage Amplifiers

Practical transistor amplifiers usually consist of a number of stages connected in cascade. In addition to providing gain, the first (or input) stage is usually required to provide a high input resistance in order to avoid loss of signal level when the amplifier is fed from a high-resistance source. In a differential amplifier the input stage must also provide large common-mode rejection. The function of the middle stages of an amplifier cascade is to provide the bulk of the voltage gain. In addition, the middle stages provide such other functions as the conversion of the signal from differential mode to single-ended mode (unless, of course, the amplifier output also is differential) and the shifting of the dc level of the signal in order to allow the output signal to swing both positive and negative. These two functions and others will be illustrated later in this section and in greater detail in Chapter 13.

Finally, the main function of the last (or output) stage of an amplifier is to provide a low output resistance in order to avoid loss of gain when a low-valued load resistance is connected to the amplifier. Also, the output stage should be able to supply the current required by the load in an efficient manner—that is, without dissipating an unduly large amount of power in the output transistors. We have already studied one type of amplifier configuration suitable for implementing output stages, namely, the source follower and the emitter follower. It will be shown in Chapter 12 that the source and emitter followers are not optimum from the point of view of power efficiency and that other, more appropriate circuit configurations exist for output stages that are required to supply large amounts of output power. In fact, we will encounter some such output stages in the op-amp circuit examples studied in Chapter 13.

To illustrate the circuit structure and the method of analysis of multistage amplifiers, we will present two examples: a two-stage CMOS op amp and a four-stage bipolar op amp.

THE INTERNATIONAL SOLID-STATE CIRCUITS CONFERENCE (ISSCC):

The year 2013 marked the sixtieth appearance of the International Solid-State Circuits Conference (ISSCC), which has become the foremost global forum for the presentation of novel developments in the broad field of solid-state circuits and systems. Sponsored by the Institute of Electrical and Electronics Engineers, ISSCC began in 1954 at the University of Pennsylvania in Philadelphia and eventually migrated to Silicon Valley with meetings annually in San Francisco since 1990.

In the early 1950s, shortly after the invention of the transistor, many of the large vacuum-tube-dependent electronics companies (e.g., GE, RCA) realized the potentially dramatic importance of these new amplifying devices. But eager as they were to create new transistor-based designs, they were quickly stymied by the lack of qualified engineers with appropriate background. In fact, Bell Labs was the only group with experience. Thus, a novel solution was needed, one in which limited expertise can be shared among multiple interested research and development entities. Thus, the first “Transistor Conference,” the precursor of ISSCC, was born. Clearly, this was an idea that worked!

9.6.1 A Two-Stage CMOS Op Amp

Figure 9.40 shows a popular structure for CMOS op amps known as the **two-stage configuration**. The circuit utilizes two power supplies, which can range from ±2.5 V for

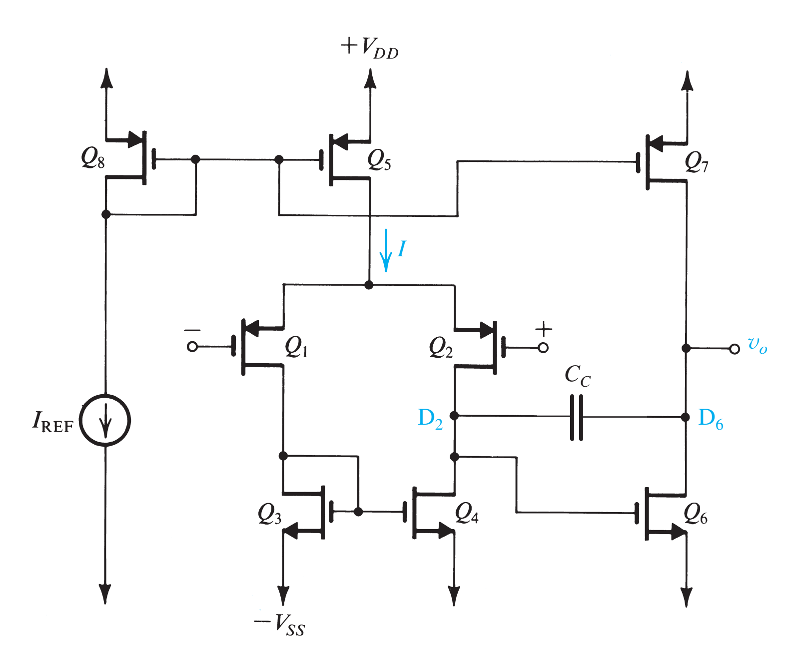


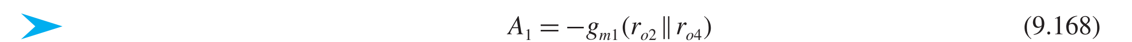
Figure 9.40 Two-stage CMOS op-amp configuration.

the 0.5-μm technology down to ±0.5 V for the 65-nm technology. A reference bias current I_{REF} is generated either externally or using on-chip circuits. One such circuit will be discussed in Chapter 13. The current mirror formed by Q_8 and Q_5 supplies the differential pair Q_1–Q_2 with bias current. The W/L ratio of Q_5 is selected to yield the desired value for the input-stage bias current I (or $I/2$ for each of Q_1 and Q_2). The input differential pair is actively loaded with the current mirror formed by Q_3 and Q_4. Thus the input stage is identical to that studied in Section 9.5 (except that here the differential pair is implemented with PMOS transistors and the current mirror with NMOS).

The second stage consists of Q_6, which is a common-source amplifier loaded with the current-source transistor Q_7. A capacitor C_C is included in the negative-feedback path of the second stage. Its function will be explained in Chapter 13, when we study the frequency response of this amplifier.

A striking feature of the circuit in Fig. 9.40 is that it does *not* have a low-output-resistance stage. In fact, the output resistance of the circuit is equal to $(r_{o6} \| r_{o7})$ and is thus rather high. This circuit, therefore, is not suitable for driving low-impedance loads. Nevertheless, the circuit is very popular and is used frequently for implementing op amps in VLSI circuits, where the op amp needs to drive only a small capacitive load, for example, in switched-capacitor circuits (Chapter 17). The simplicity of the circuit results in an op amp of reasonably good quality realized in a very small chip area.

Voltage Gain The voltage gain of the first stage was found in Section 9.5 to be given by



$$A_1 = -g_{m1}(r_{o2} \| r_{o4}) \tag{9.168}$$

where g_{m1} is the transconductance of each of the transistors of the first stage, that is, Q_1 and Q_2.

The second stage is a current-source-loaded, common-source amplifier whose voltage gain is given by

$$A_2 = -g_{m6}(r_{o6} \| r_{o7}) \tag{9.169}$$

The dc open-loop gain of the op amp is the product of A_1 and A_2.

Example 9.6

Consider the circuit in Fig. 9.40 with the following device geometries (in μm).

Transistor	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8
W/L	20/0.8	20/0.8	5/0.8	5/0.8	40/0.8	10/0.8	40/0.8	40/0.8

Let $I_{REF} = 90\ \mu A$, $V_{tn} = 0.7$ V, $V_{tp} = -0.8$ V, $\mu_n C_{ox} = 160\ \mu A/V^2$, $\mu_p C_{ox} = 40\ \mu A/V^2$, $|V_A|$ (for all devices) = 10 V, $V_{DD} = V_{SS} = 2.5$ V. For all devices, evaluate I_D, $|V_{OV}|$, $|V_{GS}|$, g_m, and r_o. Also find A_1, A_2, the dc open-loop voltage gain, the input common-mode range, and the output voltage range. Neglect the effect of V_A on bias current.

Solution

Refer to Fig. 9.40. Since Q_8 and Q_5 are matched, $I = I_{REF}$. Thus Q_1, Q_2, Q_3, and Q_4 each conducts a current equal to $I/2 = 45\ \mu A$. Since Q_7 is matched to Q_5 and Q_8, the current in Q_7 is equal to $I_{REF} = 90\ \mu A$. Finally, Q_6 conducts an equal current of $90\ \mu A$.

With I_D of each device known, we use

$$I_D = \frac{1}{2}(\mu C_{ox})(W/L)V_{OV}^2$$

to determine $|V_{OV}|$ for each transistor. Then we find $|V_{GS}|$ from $|V_{GS}| = |V_t| + |V_{OV}|$. The results are given in Table 9.1.

The transconductance of each device is determined from

$$g_m = 2I_D / |V_{OV}|$$

The value of r_o is determined from

$$r_o = |V_A| / I_D$$

The resulting values of g_m and r_o are given in Table 9.1.

Example 9.6 *continued*

Table 9.1

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8
I_D (μA)	45	45	45	45	90	90	90	90
$\|V_{OV}\|$(V)	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
$\|V_{GS}\|$(V)	1.1	1.1	1	1	1.1	1	1.1	1.1
g_m (mA/V)	0.3	0.3	0.3	0.3	0.6	0.6	0.6	0.6
r_o (kΩ)	222	222	222	222	111	111	111	111

The voltage gain of the first stage is determined from

$$A_1 = -g_{m1}(r_{o2} \| r_{o4})$$
$$= -0.3(222 \| 222) = -33.3 \text{ V/V}$$

The voltage gain of the second stage is determined from

$$A_2 = -g_{m6}(r_{o6} \| r_{o7})$$
$$= -0.6(111 \| 111) = -33.3\text{V/V}$$

Thus the overall dc open-loop gain is

$$A_0 = A_1 A_2 = (-33.3) \times (-33.3) = 1109 \text{ V/V}$$

or

$$20 \log 1109 = 61 \text{ dB}$$

The lower limit of the input common-mode range is the value of input voltage at which Q_1 and Q_2 leave the saturation region. This occurs when the input voltage falls below the voltage at the drain of Q_1 by $|V_{tp}|$ volts. Since the drain of Q_1 is at $-2.5 + 1 = -1.5$ V, then the lower limit of the input common-mode range is -2.3 V.

The upper limit of the input common-mode range is the value of input voltage at which Q_5 leaves the saturation region. Since for Q_5 to operate in saturation the voltage across it (i.e., V_{SD5}) should at least be equal to the overdrive voltage at which it is operating (i.e., 0.3 V), the highest voltage permitted at the drain of Q_5 should be $+2.2$ V. It follows that the highest value of v_{ICM} should be

$$v_{ICM\max} = 2.2 - 1.1 = 1.1\text{V}$$

The highest allowable output voltage is the value at which Q_7 leaves the saturation region, which is $V_{DD} - |V_{OV7}| = 2.5 - 0.3 = 2.2$ V. The lowest allowable output voltage is the value at which Q_6 leaves saturation, which is $-V_{SS} + V_{OV6} = -2.5 + 0.3 = -2.2$ V. Thus, the output voltage range is -2.2 V to $+2.2$ V.

Input Offset Voltage The device mismatches inevitably present in the input stage give rise to an input offset voltage. The components of this input offset voltage can be calculated using the methods developed in Section 9.4.1. Because device mismatches are random, the resulting offset voltage is referred to as **random offset**. This is to distinguish it from another type of input offset voltage that can be present even if all appropriate devices are perfectly matched. This predictable or **systematic offset** can be minimized by careful design. Although it occurs also in BJT op amps, and we have encountered it in Section 9.5.4, it is usually much more pronounced in CMOS op amps because their gain-per-stage is rather low.

To see how systematic offset can occur in the circuit of Fig. 9.40, let the two input terminals be grounded. If the input stage is perfectly balanced, then the voltage appearing at the drain of Q_4 will be equal to that at the drain of Q_3, which is $(-V_{SS}+V_{GS4})$. Now this is also the voltage that is fed to the gate of Q_6. In other words, a voltage equal to V_{GS4} appears between gate and source of Q_6. Thus the drain current of Q_6, I_6, will be related to the drain current of Q_4, which is equal to $I/2$, by the relationship

$$I_6 = \frac{(W/L)_6}{(W/L)_4}(I/2) \tag{9.170}$$

In order for no offset voltage to appear at the output, this current must be exactly equal to the current supplied by Q_7. The latter current is related to the current I of the parallel transistor Q_5 by

$$I_7 = \frac{(W/L)_7}{(W/L)_5}I \tag{9.171}$$

Now, the condition for making $I_6 = I_7$ can be found from Eqs. (9.170) and (9.171) as

$$\frac{(W/L)_6}{(W/L)_4} = 2\frac{(W/L)_7}{(W/L)_5} \tag{9.172}$$

If this condition is not met, a systematic offset will result. From the specification of the device geometries in Example 9.6, we can verify that condition (9.172) is satisfied, and, therefore, the op amp analyzed in that example should not exhibit a systematic input offset voltage.

EXERCISE

9.22 Consider the CMOS op amp of Fig. 9.40 when fabricated in a 0.8-µm CMOS technology for which $\mu_n C_{ox} = 3\mu_p C_{ox} = 90\ \mu\text{A/V}^2$, $|V_t| = 0.8$ V, and $V_{DD} = V_{SS} = 2.5$ V. For a particular design, $I = 100\ \mu$A, $(W/L)_1 = (W/L)_2 = (W/L)_5 = 200$, and $(W/L)_3 = (W/L)_4 = 100$.
(a) Find the (W/L) ratios of Q_6 and Q_7 so that $I_6 = 100\ \mu$A.
(b) Find the overdrive voltage, $|V_{OV}|$, at which each of Q_1, Q_2, and Q_6 is operating.
(c) Find g_m for Q_1, Q_2, and Q_6.
(d) If $|V_A| = 10$ V, find r_{o2}, r_{o4}, r_{o6}, and r_{o7}.
(e) Find the voltage gains A_1 and A_2, and the overall gain A.
Ans. (a) $(W/L)_6 = (W/L)_7 = 200$; (b) 0.129 V, 0.129 V, 0.105 V; (c) 0.775 mA/V, 0.775 mA/V, 1.90 mA/V; (d) 200 kΩ, 200 kΩ, 100 kΩ, 100 kΩ; (e) –77.5 V/V, –95 V/V, 7363 V/V

9.6.2 A Bipolar Op Amp

Our second example of multistage amplifiers is the four-stage bipolar op amp shown in Fig. 9.41. The circuit consists of four stages. The **differential-in**, **differential-out** input stage consists of transistors Q_1 and Q_2, which are biased by current source Q_3. The second stage is also a differential-input amplifier, but its output is taken single-endedly at the collector of Q_5. This stage is formed by Q_4 and Q_5, which are biased by the current source Q_6. Note that the conversion from differential to single-ended as performed by the second stage results in a loss of gain by a factor of 2. In the more elaborate method for accomplishing this conversion studied in Section 9.5, a current mirror was used as an active load.

In addition to providing some voltage gain, the third stage, consisting of the *pnp* transistor Q_7, provides the essential function of *shifting the dc level* of the signal. Thus, while the signal at the collector of Q_5 is not allowed to swing below the voltage at the base of Q_5 (+10 V), the signal at the collector of Q_7 can swing negatively (and positively, of course). From our study of op amps in Chapter 2, we know that the output terminal of the op amp should be capable of both positive and negative voltage swings. Therefore every op-amp circuit includes a **level-shifting** arrangement. Although the use of the complementary *pnp* transistor provides a simple solution to the level-shifting problem, other forms of level shifter exist, one of which

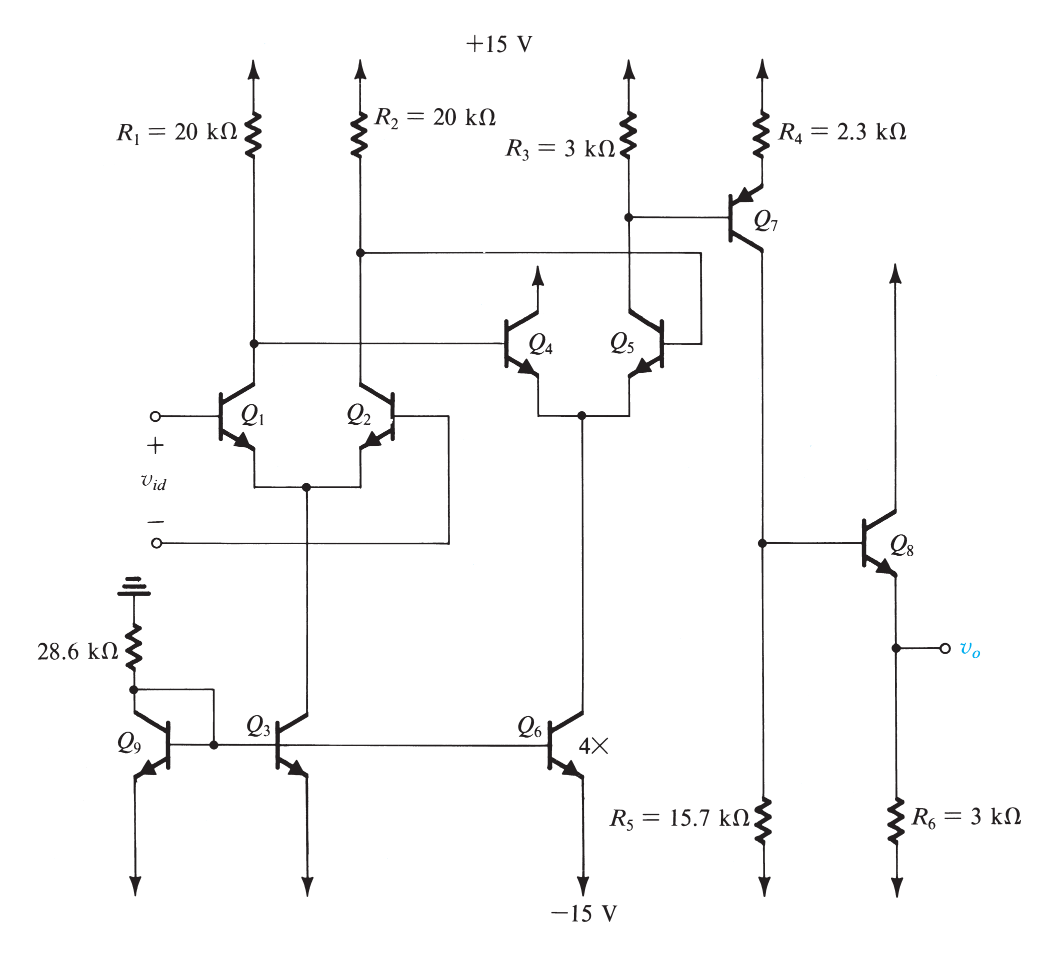


Figure 9.41 A four-stage bipolar op amp.

will be discussed in Chapter 13. Furthermore, note that level shifting is accomplished in the CMOS op amp we have studied in Section 9.6.1 by using complementary devices for the two stages: that is, *p*-channel for the first stage and *n*-channel for the second stage.

The output stage of the op amp consists of emitter follower Q_8. As we know from our study of op amps in Chapter 2, ideally the output operates around zero volts. This and other features of the BJT op amp will be illustrated in Example 9.7.

Example 9.7

In this example, we analyze the dc bias of the bipolar op-amp circuit of Fig. 9.41. Toward that end, Fig. 9.42 shows the circuit with the two input terminals connected to ground.

(a) Perform an approximate dc analysis (assuming $\beta \gg 1$, $|V_{BE}| \simeq 0.7$ V, and neglecting the Early effect) to calculate the dc currents and voltages everywhere in the circuit. Note that Q_6 has four times the area of each of Q_9 and Q_3.
(b) Calculate the quiescent power dissipation in this circuit.
(c) If transistors Q_1 and Q_2 have $\beta = 100$, calculate the input bias current of the op amp.
(d) What is the input common-mode range of this op amp?

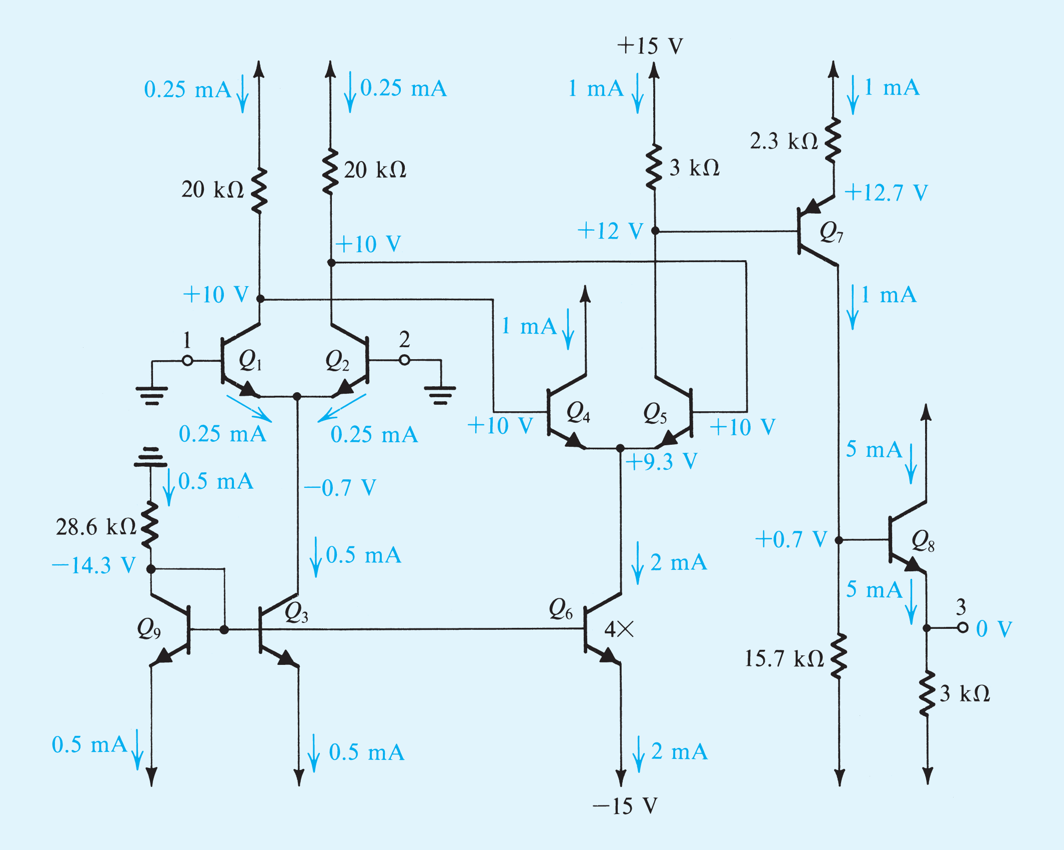


Figure 9.42 Circuit for Example 9.7.

Example 9.7 *continued*

Solution

(a) The values of all dc currents and voltages are indicated on the circuit diagram. These values were calculated by ignoring the base current of every transistor—that is, by assuming β to be very high. The analysis starts by determining the current through the diode-connected transistor Q_9 to be 0.5 mA. Then we see that transistor Q_3 conducts 0.5 mA and transistor Q_6 conducts 2 mA. The current-source transistor Q_3 feeds the differential pair (Q_1, Q_2) with 0.5 mA. Thus each of Q_1 and Q_2 will be biased at 0.25 mA. The collectors of Q_1 and Q_2 will be at $[+15 - 0.25 \times 20] = +10$ V.

Proceeding to the second differential stage formed by Q_4 and Q_5, we find the voltage at their emitters to be $[+10 - 0.7] = 9.3$ V. This differential pair is biased by the current-source transistor Q_6, which supplies a current of 2 mA; thus Q_4 and Q_5 will each be biased at 1 mA. We can now calculate the voltage at the collector of Q_5 as $[+15 - 1 \times 3] = +12$ V. This will cause the voltage at the emitter of the *pnp* transistor Q_7 to be +12.7 V, and the emitter current of Q_7 will be $(+15 - 12.7)/2.3 = 1$ mA.

The collector current of Q_7, 1 mA, causes the voltage at the collector to be $[-15 + 1 \times 15.7] = +0.7$ V. The emitter of Q_8 will be 0.7 V below the base; thus output terminal 3 will be at 0 V. Finally, the emitter current of Q_8 can be calculated to be $[0 - (-15)]/3 = 5$ mA.

(b) To calculate the power dissipated in the circuit in the quiescent state (i.e., with zero input signal) we simply evaluate the dc current that the circuit draws from each of the two power supplies. From the +15 V supply the dc current is $I^+ = 0.25 + 0.25 + 1 + 1 + 1 + 5 = 8.5$ mA. Thus the power supplied by the positive power supply is $P^+ = 15 \times 8.5 = 127.5$ mW. The –15-V supply provides a current I^- given by $I^- = 0.5 + 0.5 + 2 + 1 + 5 = 9$ mA. Thus the power provided by the negative supply is $P^- = 15 \times 9 = 135$ mW. Adding P^+ and P^- provides the total power dissipated in the circuit P_D: $P_D = P^+ + P^- = 262.5$ mW.

(c) The input bias current of the op amp is the average of the dc currents that flow in the two input terminals (i.e., in the bases of Q_1 and Q_2). These two currents are equal (because we have assumed matched devices); thus the bias current is given by

$$I_B = \frac{I_{E1}}{\beta + 1} \simeq 2.5\ \mu\text{A}$$

(d) The upper limit on the input common-mode voltage is determined by the voltage at which Q_1 and Q_2 leave the active mode and enter saturation. This will happen if the input voltage exceeds the collector voltage, which is +10 V, by about 0.4 V. Thus the upper limit of the common-mode range is +10.4 V.

The lower limit of the input common-mode range is determined by the voltage at which Q_3 leaves the active mode and thus ceases to act as a constant-current source. This will happen if the collector voltage of Q_3 goes below the voltage at its base, which is –14.3 V, by more than 0.4 V. It follows that the input common-mode voltage should not go lower than $-14.7 + 0.7 = -14$ V. Thus the common-mode range is –14 V to +10.4 V.

Example 9.8

Use the dc bias quantities evaluated in Example 9.7 to analyze the circuit in Fig. 9.41 to determine the input resistance, the voltage gain, and the output resistance.

Solution

The input differential resistance R_{id} is given by

$$R_{id} = r_{\pi 1} + r_{\pi 2}$$

Since each of Q_1 and Q_2 is operating at an emitter current of 0.25 mA, it follows that

$$r_{e1} = r_{e2} = \frac{25}{0.25} = 100\ \Omega$$

Assume $\beta = 100$; then

$$r_{\pi 1} = r_{\pi 2} = 101 \times 100 = 10.1\ \text{k}\Omega$$

Thus,

$$R_{id} = 20.2\ \text{k}\Omega$$

To evaluate the gain of the first stage, we first find the input resistance of the second stage, R_{i2},

$$R_{i2} = r_{\pi 4} + r_{\pi 5}$$

Q_4 and Q_5 are each operating at an emitter current of 1 mA; thus

$$r_{e4} = r_{e5} = 25\ \Omega$$
$$r_{\pi 4} = r_{\pi 5} = 101 \times 25 = 2.525\ \text{k}\Omega$$

Thus $R_{i2} = 5.05\ \text{k}\Omega$. This resistance appears between the collectors of Q_1 and Q_2, as shown in Fig. 9.43. Thus the gain of the first stage will be

$$A_1 \equiv \frac{v_{o1}}{v_{id}} \simeq \frac{\text{Total resistance in collector circuit}}{\text{Total resistance in emitter circuit}}$$

$$= \frac{R_{i2} \parallel (R_1 + R_2)}{r_{e1} + r_{e2}}$$

$$= \frac{5.05\ \text{k}\Omega \parallel 40\ \text{k}\Omega}{200\ \Omega} = 22.4\ \text{V/V}$$

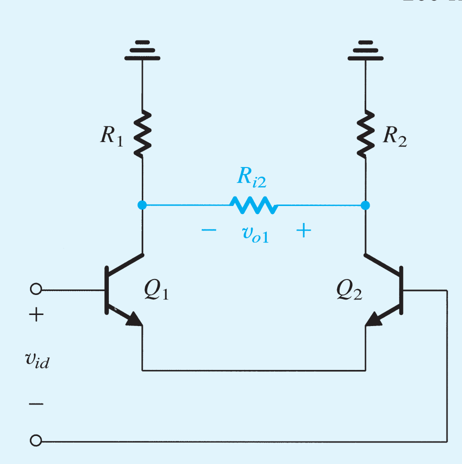


Figure 9.43 Equivalent circuit for calculating the gain of the input stage of the amplifier in Fig. 9.41.

Example 9.8 *continued*

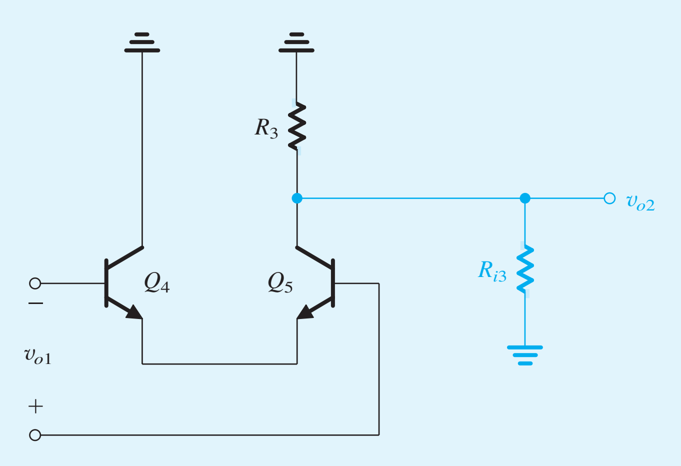


Figure 9.44 Equivalent circuit for calculating the gain of the second stage of the amplifier in Fig. 9.41.

Figure 9.44 shows an equivalent circuit for calculating the gain of the second stage. As indicated, the input voltage to the second stage is the output voltage of the first stage, v_{o1}. Also shown is the resistance R_{i3}, which is the input resistance of the third stage formed by Q_7. The value of R_{i3} can be found by multiplying the total resistance in the emitter of Q_7 by $(\beta+1)$:

$$R_{i3} = (\beta+1)(R_4 + r_{e7})$$

Since Q_7 is operating at an emitter current of 1 mA,

$$r_{e7} = \frac{25}{1} = 25\ \Omega$$

$$R_{i3} = 101 \times 2.325 = 234.8\ \text{k}\Omega$$

We can now find the gain A_2 of the second stage as the ratio of the total resistance in the collector circuit to the total resistance in the emitter circuit:

$$A_2 \equiv \frac{v_{o2}}{v_{o1}} \simeq -\frac{R_3 \| R_{i3}}{r_{e4} + r_{e5}}$$

$$= -\frac{3\ \text{k}\Omega \| 234.8\ \text{k}\Omega}{50\ \Omega} = -59.2\ \text{V/V}$$

To obtain the gain of the third stage we refer to the equivalent circuit shown in Fig. 9.45, where R_{i4} is the input resistance of the output stage formed by Q_8. Using the resistance-reflection rule, we calculate the value of R_{i4} as

$$R_{i4} = (\beta+1)(r_{e8} + R_6)$$

where

$$r_{e8} = \frac{25}{5} = 5\ \Omega$$

$$R_{i4} = 101(5 + 3000) = 303.5\ \text{k}\Omega$$

The gain of the third stage is given by

$$A_3 \equiv \frac{v_{o3}}{v_{o2}} \simeq -\frac{R_5 \| R_{i4}}{r_{e7} + R_4}$$

$$= -\frac{15.7\ \text{k}\Omega \| 303.5\ \text{k}\Omega}{2.325\ \text{k}\Omega} = -6.42\ \text{V/V}$$

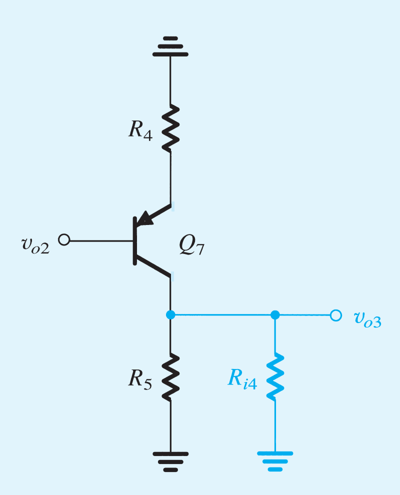


Figure 9.45 Equivalent circuit for evaluating the gain of the third stage in the amplifier circuit of Fig. 9.41.

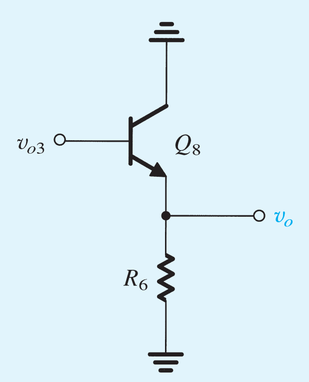


Figure 9.46 Equivalent circuit of the output stage of the amplifier circuit of Fig. 9.41.

Example 9.8 *continued*

Finally, to obtain the gain A_4 of the output stage we refer to the equivalent circuit in Fig. 9.46 and write

$$A_4 \equiv \frac{v_o}{v_{o3}} = \frac{R_6}{R_6 + r_{e8}}$$

$$= \frac{3000}{3000 + 5} = 0.998 \simeq 1$$

The overall voltage gain of the amplifier can then be obtained as follows:

$$\frac{v_o}{v_{id}} = A_1 A_2 A_3 A_4 = 8513 \text{ V/V}$$

or 78.6 dB.

To obtain the output resistance R_o we "grab hold" of the output terminal in Fig. 9.41 and look back into the circuit. By inspection we find

$$R_o = R_6 \parallel [r_{e8} + R_5/(\beta + 1)]$$

which gives

$$R_o = 152 \ \Omega$$

EXERCISE

9.23 Use the results of Example 9.8 to calculate the overall voltage gain of the amplifier in Fig. 9.41 when it is connected to a source having a resistance of 10 kΩ and a load of 1 kΩ.
Ans. 4943 V/V

Analysis Using Current Gains There is an alternative method for the analysis of bipolar multistage amplifiers that can be somewhat easier to perform in some cases. The method makes use of current gains or more appropriately current-transmission factors. In effect, one traces the transmission of the signal current throughout the amplifier cascade, evaluating all the current transmission factors in turn. We shall illustrate the method by using it to analyze the amplifier circuit of the preceding example.

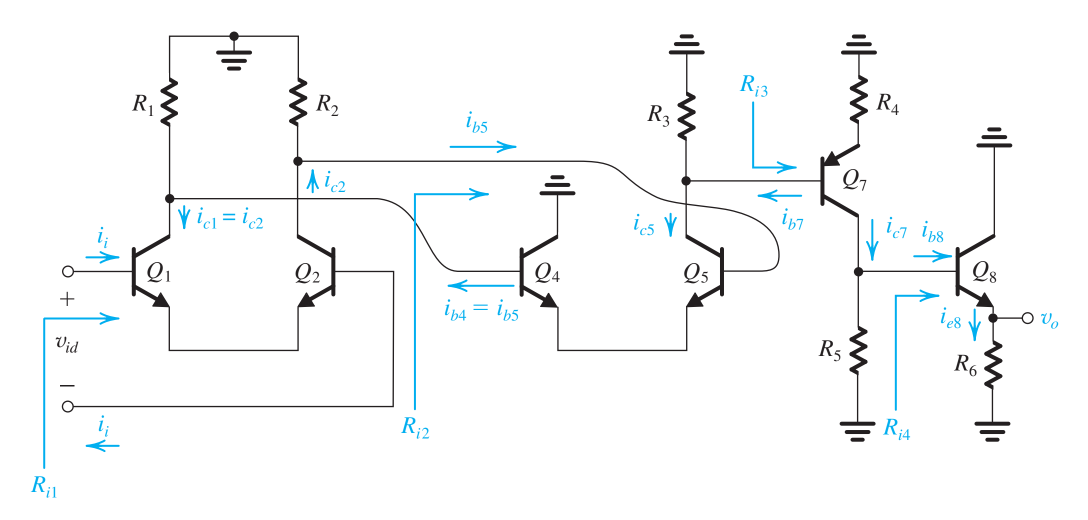


Figure 9.47 The circuit of the multistage amplifier of Fig. 9.41 prepared for small-signal analysis. Indicated are the signal currents throughout the amplifier and the input resistances of the four stages.

Figure 9.47 shows the amplifier circuit prepared for small-signal analysis. We have indicated on the circuit diagram the signal currents through all the circuit branches. Also indicated are the input resistances of all four stages of the amplifier. These should be evaluated before commencing the following analysis.

The purpose of the analysis is to determine the overall voltage gain (v_o/v_{id}). Toward that end, we express v_o in terms of the signal current in the emitter of Q_8, i_{e8}, and v_{id} in terms of the input signal current i_i, as follows:

$$v_o = R_6 i_{e8}$$
$$v_{id} = R_{i1} i_i$$

Thus, the voltage gain can be expressed in terms of the current gain (i_{e8}/i_i) as

$$\frac{v_o}{v_{id}} = \frac{R_6}{R_{i1}} \frac{i_{e8}}{i_i}$$

Next, we expand the current gain (i_{e8}/i_i) in terms of the signal currents throughout the circuit as follows:

$$\frac{i_{e8}}{i_i} = \frac{i_{e8}}{i_{b8}} \times \frac{i_{b8}}{i_{c7}} \times \frac{i_{c7}}{i_{b7}} \times \frac{i_{b7}}{i_{c5}} \times \frac{i_{c5}}{i_{b5}} \times \frac{i_{b5}}{i_{c2}} \times \frac{i_{c2}}{i_i}$$

Each of the current-transmission factors on the right-hand side is either the current gain of a transistor or the ratio of a current divider. Thus, reference to Fig. 9.47 enables us to find these

factors by inspection:

$$\frac{i_{e8}}{i_{b8}} = \beta_8 + 1$$

$$\frac{i_{b8}}{i_{c7}} = \frac{R_5}{R_5 + R_{i4}}$$

$$\frac{i_{c7}}{i_{b7}} = \beta_7$$

$$\frac{i_{b7}}{i_{c5}} = \frac{R_3}{R_3 + R_{i3}}$$

$$\frac{i_{c5}}{i_{b5}} = \beta_5$$

$$\frac{i_{b5}}{i_{c2}} = \frac{(R_1 + R_2)}{(R_1 + R_2) + R_{i2}}$$

$$\frac{i_{c2}}{i_i} = \beta_2$$

These ratios can be easily evaluated and their values used to determine the voltage gain.

With a little practice, it is possible to carry out such an analysis very quickly, foregoing explicitly labeling the signal currents on the circuit diagram. One simply "walks through" the circuit, from input to output, or vice versa, determining the current-transmission factors one at a time, in a chainlike fashion.

EXERCISE

9.24 Use the values of input resistance found in Example 9.8 to evaluate the seven current-transmission factors and hence the overall current gain and voltage gain.
Ans. The current-transmission factors in the order of their listing are 101, 0.0492, 100, 0.0126, 100, 0.8879, 100 A/A; the overall current gain is 55599 A/A; the voltage gain is 8257 V/V. This value differs slightly from that found in Example 9.8, because of the various approximations made in the example (e.g., $\alpha \simeq 1$).

Summary

- The differential-pair or differential-amplifier configuration is the most widely used building block in analog IC design. The input stage of every op amp is a differential amplifier.

- There are two reasons for preferring differential to single-ended amplifiers: Differential amplifiers are insensitive to interference, and they do not need bypass and coupling capacitors.

- For a MOS (bipolar) pair biased by a current source I, each device operates at a drain (collector, assuming $\alpha = 1$) current of $I/2$ and a corresponding overdrive voltage V_{OV} (no counterpart in bipolar). Each device has $g_m = I/V_{OV}$ ($\alpha I/2V_T$, for bipolar) and $r_o = |V_A|/(I/2)$.

- With the two input terminals connected to a suitable dc voltage V_{CM}, the bias current I of a perfectly symmetrical differential pair divides equally between the

two transistors of the pair, resulting in a zero voltage difference between the two drains (collectors). To steer the current completely to one side of the pair, a difference input voltage v_{id} of at least $\sqrt{2}V_{OV}$ ($4V_T$ for bipolar) is needed.

- Superimposing a differential input signal v_{id} on the dc common-mode input voltage V_{CM} such that $v_{I1} = V_{CM} + v_{id}/2$ and $v_{I2} = V_{CM} - v_{id}/2$ causes a virtual signal ground to appear on the common-source (common-emitter) connection. In response to v_{id}, the current in Q_1 increases by $g_m v_{id}/2$ and the current in Q_2 decreases by $g_m v_{id}/2$. Thus, voltage signals of $\pm g_m(R_D \| r_o)v_{id}/2$ develop at the two drains (collectors, with R_D replaced by R_C). If the output voltage is taken single-endedly, that is, between one of the drains (collectors) and ground, a differential gain of $\frac{1}{2}g_m(R_D \| r_o)$ is realized. When the output is taken differentially, that is, between the two drains (collectors), the differential gain realized is twice as large: $g_m(R_D \| r_o)$.
- The analysis of a differential amplifier to determine differential gain, differential input resistance, frequency response of differential gain, and so on is facilitated by employing the differential half-circuit, which is a common-source (common-emitter) transistor biased at $I/2$.
- An input common-mode signal v_{icm} gives rise to drain (collector) voltage signals that are ideally equal and given by $-v_{icm}(R_D/2R_{SS})[-v_{icm}(R_C/2R_{EE})$ for the bipolar pair], where R_{SS} (R_{EE}) is the output resistance of the current source that supplies the bias current I. When the output is taken single-endedly, a common-mode gain of magnitude $|A_{cm}| = R_D/2R_{SS}$ ($R_C/2R_{EE}$ for the bipolar case) results. Taking the output differentially results, in the perfectly matched case, in zero A_{cm} (infinite CMRR). Mismatches between the two sides of the pair make A_{cm} finite even when the output is taken differentially: A mismatch ΔR_D causes $|A_{cm}| = (R_D/2R_{SS})(\Delta R_D/R_D)$; a mismatch Δg_m causes $|A_{cm}| = (R_D/2R_{SS})$ $(\Delta g_m/g_m)$. Corresponding expressions apply for the bipolar pair.
- While the input differential resistance R_{id} of the MOS pair is infinite, that for the bipolar pair is only $2r_\pi$ but can be increased to $2(\beta+1)(r_e + R_e)$ by including resistances R_e in the two emitters. The latter action, however, lowers A_d.
- Mismatches between the two sides of a differential pair result in a differential dc output voltage V_O even when the two input terminals are tied together and connected to a dc voltage V_{CM}. This signifies the presence of an input offset voltage $V_{OS} \equiv V_O/A_d$. In a MOS pair there are three main sources for V_{OS}:

$$\Delta R_D \Rightarrow V_{OS} = \frac{V_{OV}}{2}\frac{\Delta R_D}{R_D}$$

$$\Delta(W/L) \Rightarrow V_{OS} = \frac{V_{OV}}{2}\frac{\Delta(W/L)}{W/L}$$

$$\Delta V_t \Rightarrow V_{OS} = \Delta V_t$$

For the bipolar pair there are two main sources:

$$\Delta R_C \Rightarrow V_{OS} = V_T\frac{\Delta R_C}{R_C}$$

$$\Delta I_S \Rightarrow V_{OS} = V_T\frac{\Delta I_S}{I_S}$$

- A popular circuit in both MOS and bipolar analog ICs is the current-mirror-loaded differential pair. It realizes a high differential gain $A_d = g_m(R_{o\ \text{pair}} \| R_{o\ \text{mirror}})$ and a low common-mode gain, $|A_{cm}| = 1/2g_{m3}R_{SS}$ for the MOS circuit ($r_{o4}/\beta_3 R_{EE}$ for the bipolar circuit), as well as performing the differential-to-single-ended conversion with no loss of gain.
- The CMOS two-stage amplifier studied in Section 9.6.1 is intended for use as part of an IC system and thus is required to drive only small capacitive loads. Therefore it does not have an output stage with a low output resistance.
- A multistage amplifier typically consists of three or more stages: an input stage having a high input resistance, a reasonably high gain, and, if differential, a high CMRR; one or two intermediate stages that realize the bulk of the gain; and an output stage having a low output resistance. In designing and analyzing a multistage amplifier, the loading effect of each stage on the one that precedes it must be taken into account.

Computer Simulation Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as allowable signal swing and amplifier nonlinear distortion. Instructions to assist in setting up PSpice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 9.1: The MOS Differential Pair

9.1 For an NMOS differential pair with a common-mode voltage V_{CM} applied, as shown in Fig. 9.2, let $V_{DD} = V_{SS} = 1.0$ V, $k_n' = 0.4$ mA/V^2, $(W/L)_{1,2} = 10$, $V_{tn} = 0.4$ V, $I = 0.16$ mA, $R_D = 5$ kΩ, and neglect channel-length modulation.

(a) Find V_{OV} and V_{GS} for each transistor.
(b) For $V_{CM} = 0$, find V_S, I_{D1}, I_{D2}, V_{D1}, and V_{D2}.
(c) Repeat (b) for $V_{CM} = +0.4$ V.
(d) Repeat (b) for $V_{CM} = -0.1$ V.
(e) What is the highest value of V_{CM} for which Q_1 and Q_2 remain in saturation?
(f) If current source I requires a minimum voltage of 0.2 V to operate properly, what is the lowest value allowed for V_S and hence for V_{CM}?

9.2 For the PMOS differential amplifier shown in Fig. P9.2 let $V_{tp} = -0.8$ V and $k_p' W/L = 4$ mA/V^2. Neglect channel-length modulation.

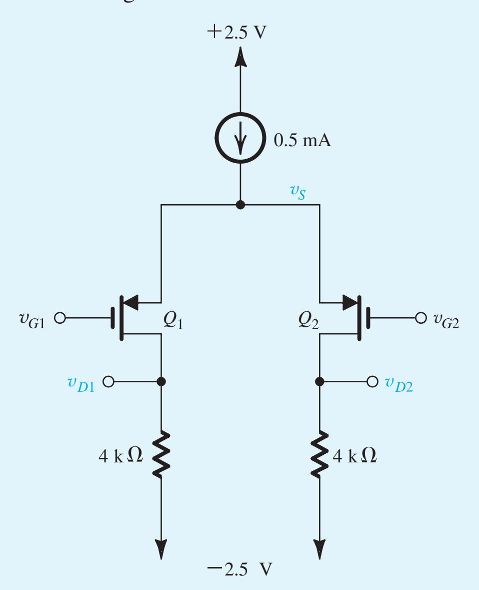


Figure P9.2

(a) For $v_{G1} = v_{G2} = 0$ V, find $|V_{OV}|$ and V_{SG} for each of Q_1 and Q_2. Also find V_S, V_{D1}, and V_{D2}.
(b) If the current source requires a minimum voltage of 0.4 V, find the input common-mode range.

9.3 For the differential amplifier specified in Problem 9.1 let $v_{G2} = 0$ and $v_{G1} = v_{id}$. Find the value of v_{id} that corresponds to each of the following situations:
(a) $i_{D1} = i_{D2} = 0.08$ mA; (b) $i_{D1} = 0.12$ mA and $i_{D2} = 0.04$ mA; (c) $i_{D1} = 0.16$ mA and $i_{D2} = 0$ (Q_2 just cuts off); (d) $i_{D1} = 0.04$ mA and $i_{D2} = 0.12$ mA; (e) $i_{D1} = 0$ mA (Q_1 just cuts off) and $i_{D2} = 0.16$ mA. For each case, find v_S, v_{D1}, v_{D2}, and $(v_{D2} - v_{D1})$.

SIM **9.4** For the differential amplifier specified in Problem 9.2, let $v_{G2} = 0$ and $v_{G1} = v_{id}$. Find the range of v_{id} needed to steer the bias current from one side of the pair to the other. At each end of this range, give the value of the voltage at the common-source terminal and the drain voltages.

9.5 Consider the differential amplifier specified in Problem 9.1 with G_2 grounded and $v_{G1} = v_{id}$. Let v_{id} be adjusted to the value that causes $i_{D1} = 0.09$ mA and $i_{D2} = 0.07$ mA. Find the corresponding values of v_{GS2}, v_S, v_{GS1}, and hence v_{id}. What is the difference output voltage $v_{D2} - v_{D1}$? What is the voltage gain $(v_{D2} - v_{D1})/v_{id}$? What value of v_{id} results in $i_{D1} = 0.07$ mA and $i_{D2} = 0.09$ mA?

D 9.6 Design the circuit in Fig. P9.6 to obtain a dc voltage of +0.1 V at each of the drains of Q_1 and Q_2 when $v_{G1} = v_{G2} = 0$ V. Operate all transistors at $V_{OV} = 0.15$ V

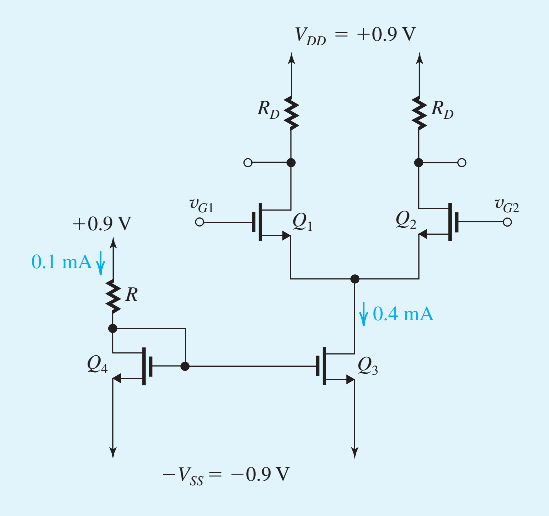


Figure P9.6

and assume that for the process technology in which the circuit is fabricated, $V_{tn} = 0.4$ V and $\mu_n C_{ox} = 400\ \mu\text{A/V}^2$. Neglect channel-length modulation. Determine the values of R, R_D, and the W/L ratios of Q_1, Q_2, Q_3, and Q_4. What is the input common-mode voltage range for your design?

9.7 The table providing the answers to Exercise 9.3 shows that as the maximum input signal to be applied to the differential pair is increased, linearity is maintained at the same level by operating at a higher V_{OV}. If $|v_{id}|_{\max}$ is to be 220 mV, use the data in the table to determine the required V_{OV} and the corresponding values of W/L and g_m.

9.8 Use Eq. (9.23) to show that if the term involving v_{id}^2 is to be kept to a maximum value of k then the maximum possible fractional change in the transistor current is given by

$$\frac{\Delta I_{\max}}{I/2} = 2\sqrt{k(1-k)}$$

and the corresponding maximum value of v_{id} is given by

$$v_{id\max} = 2\sqrt{k}\,V_{OV}$$

Evaluate both expressions for $k = 0.01$, 0.1, and 0.2.

9.9 A MOS differential amplifier biased with a current source $I = 200\ \mu$A is found to switch currents completely to one side of the pair when a difference signal $v_{id} = 0.3$ V is applied. At what overdrive voltage will each of Q_1 and Q_2 be operating when $v_{id} = 0$? If v_{id} for full current switching is to be 0.5 V, what must the bias current I be changed to?

D 9.10 Design the MOS differential amplifier of Fig. 9.5 to operate at $V_{OV} = 0.25$ V and to provide a transconductance g_m of 1 mA/V. Specify the W/L ratios and the bias current. The technology available provides $V_t = 0.5$ V and $\mu_n C_{ox} = 400\ \mu\text{A/V}^2$.

9.11 For the MOS differential pair in Fig. 9.5, specify the value of $v_{id} \equiv v_{G1} - v_{G2}$, in terms of V_{OV}, that

(a) causes i_{D1} to increase by 10% above its equilibrium value of $I/2$.
(b) makes $i_{D1}/i_{D2} = 1.0$; 2.0; 1.1; 1.01; 20.

9.12 An NMOS differential amplifier is operated at a bias current I of 0.2 mA and has a W/L ratio of 32, $\mu_n C_{ox} = 200\ \mu\text{A/V}^2$, $V_A = 10$ V, and $R_D = 10$ kΩ. Find V_{OV}, g_m, r_o, and A_d.

D 9.13 It is required to design an NMOS differential amplifier to operate with a differential input voltage that can be as high as 0.1 V while keeping the nonlinear term under the square root in Eq. (9.23) to a maximum of 0.04. A transconductance g_m of 2 mA/V is needed and the amplifier is required to provide a differential output signal of 1 V when the input is at its maximum value. Find the required values of V_{OV}, I, R_D, and W/L. Assume that the technology available has $\mu_n C_{ox} = 200\ \mu\text{A/V}^2$ and $\lambda = 0$.

D 9.14 Design a MOS differential amplifier to operate from ±1-V power supplies and dissipate no more than 1 mW in the equilibrium state. The differential voltage gain A_d is to be 10 V/V and the output common-mode dc voltage is to be 0.2 V. (*Note*: This is the dc voltage at the drains.) Assume $\mu_n C_{ox} = 400\ \mu\text{A/V}^2$ and neglect the Early effect. Specify I, R_D, and W/L.

D 9.15 Design a MOS differential amplifier to operate from ±1-V supplies and dissipate no more than 1 mW in its equilibrium state. Select the value of V_{OV} so that the value of v_{id} that steers the current from one side of the pair to the other is 0.25 V. The differential voltage gain A_d is to be 10 V/V. Assume $k_n' = 400\ \mu\text{A/V}^2$ and neglect the Early effect. Specify the required values of I, R_D, and W/L.

9.16 An NMOS differential amplifier employing equal drain resistors, $R_D = 47$ kΩ, has a differential gain A_d of 20 V/V.

(a) What is the value of g_m for each of the two transistors?
(b) If each of the two transistors is operating at an overdrive voltage $V_{OV} = 0.2$ V, what must the value of I be?
(c) For $v_{id} = 0$, what is the dc voltage across each R_D?
(d) If v_{id} is 20-mV peak-to-peak sine wave applied in a balanced manner but superimposed on $V_{CM} = 0.5$ V, what is the peak of the sine-wave signal at each drain?
(e) What is the lowest value that V_{DD} must have to ensure saturation-mode operation for Q_1 and Q_2 at all times? Assume $V_t = 0.5$ V.

9.17 A MOS differential amplifier is designed to have a differential gain A_d equal to the voltage gain obtained from a common-source amplifier. Both amplifiers utilize the same values of R_D and supply voltages, and all the transistors have the same W/L ratios. What must the bias current I of the differential pair be relative to the bias current I_D of the CS amplifier? What is the ratio of the power dissipation of the two circuits?

9.18 A differential amplifier is designed to have a differential voltage gain equal to the voltage gain of a common-source amplifier. Both amplifiers use the same values of R_D and supply voltages and are designed to dissipate equal amounts of power in their equilibrium or quiescent state. As well, all the transistors use the same channel length. What must the width W of the differential-pair transistors be relative to the width of the CS transistor?

D 9.19 Figure P9.19 shows a MOS differential amplifier with the drain resistors R_D implemented using diode-connected PMOS transistors, Q_3 and Q_4. Let Q_1 and Q_2 be matched, and Q_3 and Q_4 be matched.

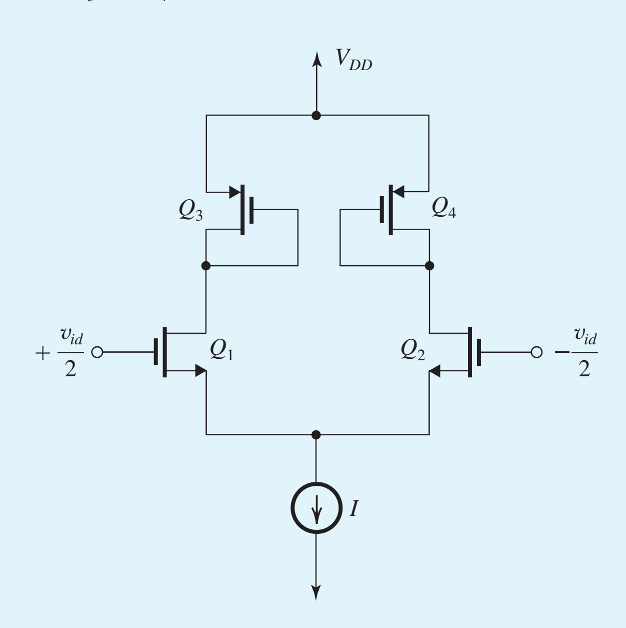


Figure P9.19

(a) Find the differential half-circuit and use it to derive an expression for A_d in terms of $g_{m1,2}$, $g_{m3,4}$, $r_{o1,2}$, and $r_{o3,4}$.
(b) Neglecting the effect of the output resistances r_o, find A_d in terms of μ_n, μ_p, $(W/L)_{1,2}$ and $(W/L)_{3,4}$.
(c) If $\mu_n = 4\mu_p$ and all four transistors have the same channel length, find $(W_{1,2}/W_{3,4})$ that results in $A_d = 10$ V/V.

9.20 Find the differential half-circuit for the differential amplifier shown in Fig. P9.20 and use it to derive an expression for the differential gain $A_d \equiv v_{od}/v_{id}$ in terms of g_m, R_D, and R_s. Neglect the Early effect. What is the gain with $R_s = 0$? What is the value of R_s (in terms of $1/g_m$) that reduces the gain to half this value?

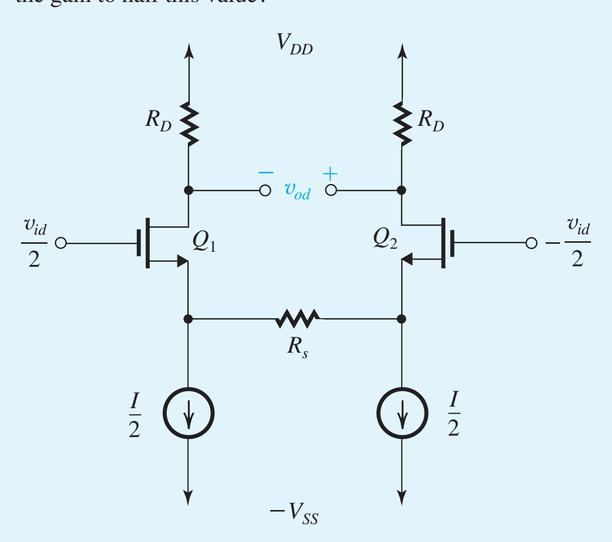


Figure P9.20

***9.21** The resistance R_s in the circuit of Fig. P9.20 can be implemented by using a MOSFET operated in the triode region, as shown in Fig. P9.21. Here Q_3 implements R_s, with the value of R_s determined by the voltage V_C at the gate of Q_3.

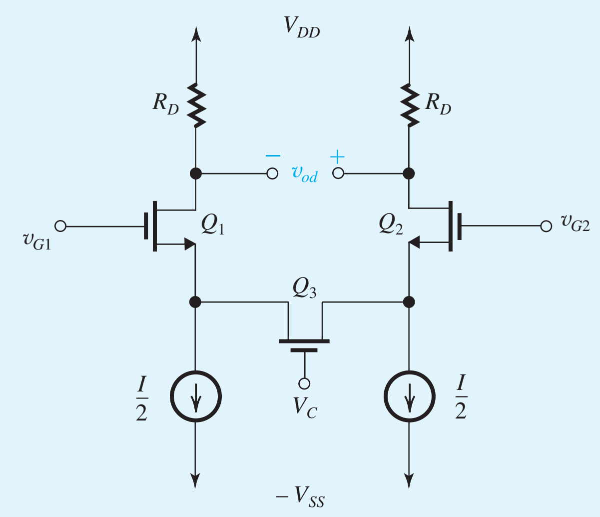


Figure P9.21

(a) With $v_{G1} = v_{G2} = 0$ V, and assuming that Q_1 and Q_2 are operating in saturation, what dc voltages appear

at the sources of Q_1 and Q_2? Express these in terms of the overdrive voltage V_{OV} at which each of Q_1 and Q_2 operates, and V_t.

(b) For the situation in (a), what current flows in Q_3? What overdrive voltage V_{OV3} is Q_3 operating at, in terms of V_C, V_{OV}, and V_t?

(c) Now consider the case $v_{G1} = +v_{id}/2$ and $v_{G2} = -v_{id}/2$, where v_{id} is a small signal. Convince yourself that Q_3 now conducts current and operates in the triode region with a small v_{DS}. What resistance r_{DS} does it have, expressed in terms of the overdrive voltage V_{OV3} at which it is operating? This is the resistance R_s. Now if all three transistors have the same W/L, express R_s in terms of V_{OV}, V_{OV3}, and $g_{m1,2}$.

(d) Find V_{OV3} and hence V_C that result in (i) $R_s = 1/g_{m1,2}$; (ii) $R_s = 0.5/g_{m1,2}$.

***9.22** The circuit of Fig. P9.22 shows an effective way of implementing the resistance R_s needed for the circuit in Fig. P9.20. Here R_s is realized as the series equivalent of two MOSFETs Q_3 and Q_4 that are operated in the triode region, thus, $R_s = r_{DS3} + r_{DS4}$. Assume that Q_1 and Q_2 are matched and operate in saturation at an overdrive voltage V_{OV} that corresponds to a drain bias current of $I/2$. Also, assume that Q_3 and Q_4 are matched.

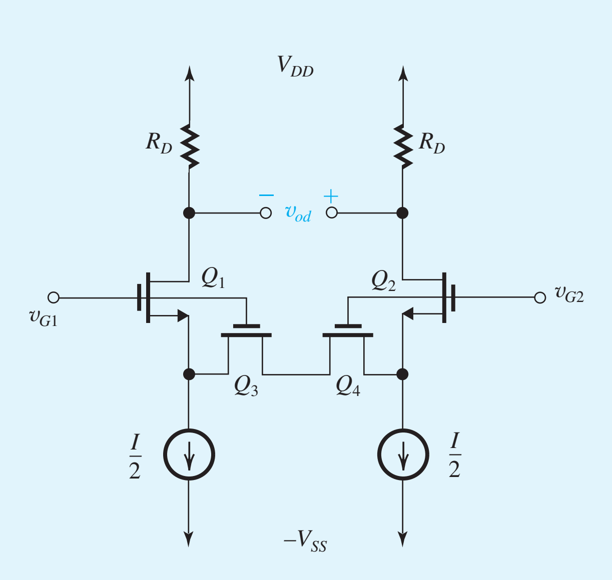


Figure P9.22

(a) With $v_{G1} = v_{G2} = 0$ V, what dc voltages appear at the sources of Q_1 and Q_2? What current flows through Q_3 and Q_4? At what overdrive voltages are Q_3 and Q_4 operating? Find an expression for r_{DS} for each of Q_3 and Q_4 and hence for R_s in terms of $(W/L)_{1,2}$, $(W/L)_{3,4}$, and $g_{m1,2}$.

(b) Now with $v_{G1} = v_{id}/2$ and $v_{G2} = -v_{id}/2$, where v_{id} is a small signal, find an expression of the voltage gain $A_d \equiv v_{od}/v_{id}$ in terms of $g_{m1,2}$, R_D, $(W/L)_{1,2}$, and $(W/L)_{3,4}$.

D *9.23 Figure P9.23 shows a circuit for a differential amplifier with an active load. Here Q_1 and Q_2 form the differential pair, while the current source transistors Q_4 and Q_5 form the active loads for Q_1 and Q_2, respectively. The dc bias circuit that establishes an appropriate dc voltage at the drains of Q_1 and Q_2 is not shown. It is required to design the circuit to meet the following specifications:

(a) Differential gain $A_d = 50$ V/V.
(b) $I_{REF} = I = 200$ μA.
(c) The dc voltage at the gates of Q_6 and Q_3 is +0.8 V.
(d) The dc voltage at the gates of Q_7, Q_4, and Q_5 is −0.8 V.

The technology available is specified as follows: $\mu_n C_{ox} = 2.5\,\mu_p C_{ox} = 250\ \mu\text{A/V}^2$; $V_{tn} = |V_{tp}| = 0.5$ V, $V_{An} = |V_{Ap}| = 10$ V. Specify the required value of R and the W/L ratios for all transistors. Also specify I_D and $|V_{GS}|$ at which each transistor is operating. For dc bias calculations you may neglect channel-length modulation.

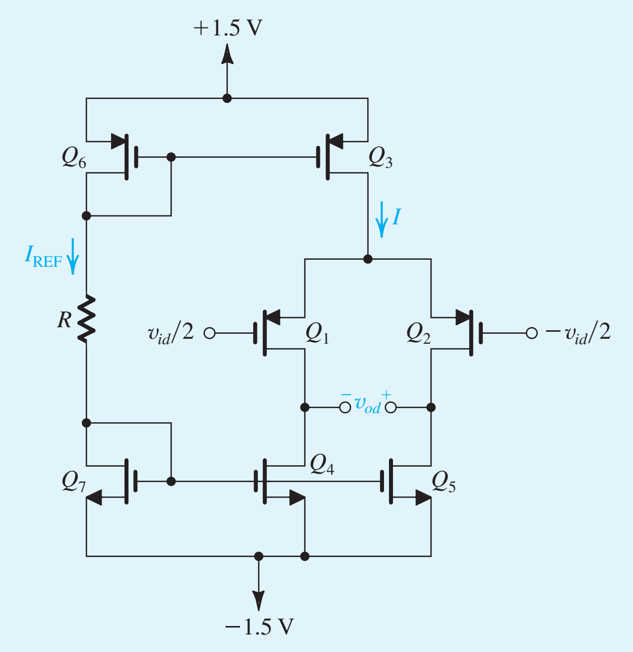


Figure P9.23

***9.24** A design error has resulted in a gross mismatch in the circuit of Fig. P9.24. Specifically, Q_2 has twice the W/L ratio of Q_1. If v_{id} is a small sine-wave signal, find:

(a) I_{D1} and I_{D2}.
(b) V_{OV} for each of Q_1 and Q_2.
(c) The differential gain A_d in terms of R_D, I, and V_{OV}.

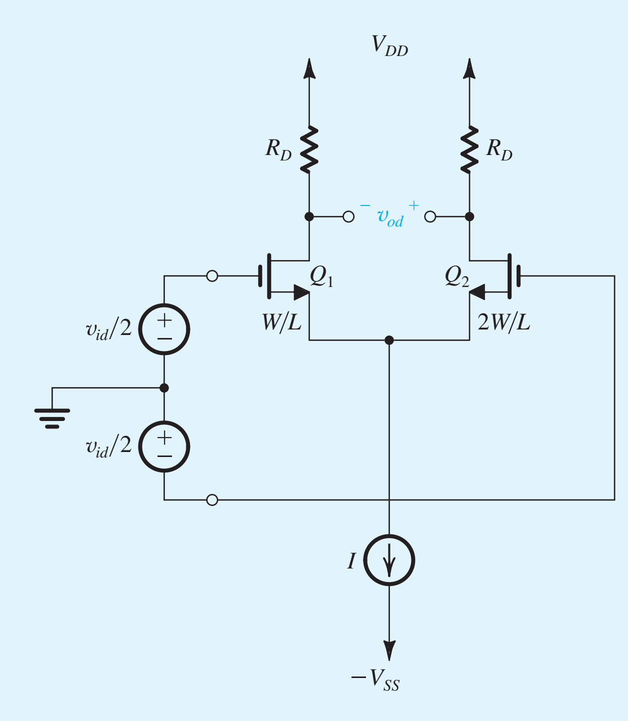


Figure P9.24

D 9.25 For the cascode differential amplifier of Fig. 9.13(a), show that if all transistors have the same channel length and are operated at the same $|V_{OV}|$ and assuming that $V'_{An} = |V'_{Ap}| = |V'_A|$, the differential gain A_d is given by

$$A_d = 2(|V_A|/|V_{OV}|)^2$$

Now design the amplifier to obtain a differential gain of 500 V/V. Use $|V_{OV}| = 0.2$ V. If $|V'_A| = 5$ V/μm, specify the required channel length L. If g_m is to be as high as possible but the power dissipation in the amplifier (in equilibrium) is to be limited to 0.5 mW, what bias current I would you use? Let $V_{DD} = V_{SS} = 0.9$ V.

Section 9.2: The BJT Differential Pair

9.26 For the differential amplifier of Fig. 9.15(a) let $I = 0.4$ mA, $V_{CC} = V_{EE} = 2.5$ V, $V_{CM} = -1$ V, $R_C = 5$ kΩ, and $\beta = 100$. Assume that the BJTs have $v_{BE} = 0.7$ V at $i_C = 1$ mA. Find the voltage at the emitters and at the outputs.

9.27 An *npn* differential amplifier with $I = 0.4$ mA, $V_{CC} = V_{EE} = 2.5$ V, and $R_C = 5$ kΩ utilizes BJTs with $\beta = 100$ and $v_{BE} = 0.7$ V at $i_C = 1$ mA. If $v_{B2} = 0$, find V_E, V_{C1}, and V_{C2} obtained with $v_{B1} = +0.5$ V, and with $v_{B1} = -0.5$ V. Assume that the current source requires a minimum of 0.3 V for proper operation.

9.28 An *npn* differential amplifier with $I = 0.4$ mA, $V_{CC} = V_{EE} = 2.5$ V, and $R_C = 5$ kΩ utilizes BJTs with $\beta = 100$ and $v_{BE} = 0.7$ V at $i_C = 1$ mA. Assuming that the bias current is obtained by a simple current source and that all transistors require a minimum v_{CE} of 0.3 V for operation in the active mode, find the input common-mode range.

9.29 Repeat Exercise 9.7 for an input of –0.3 V.

9.30 An *npn* differential pair employs transistors for which $v_{BE} = 690$ mV at $i_C = 1$ mA, and $\beta = 50$. The transistors leave the active mode at $v_{CE} \leq 0.3$ V. The collector resistors $R_C = 82$ kΩ, and the power supplies are ±1.2 V. The bias current $I = 20$ μA and is supplied with a simple current source.

(a) For $v_{B1} = v_{B2} = V_{CM} = 0$ V, find V_E, V_{C1}, and V_{C2}.
(b) Find the input common-mode range.
(c) If $v_{B2} = 0$, find the value of v_{B1} that increases the current in Q_1 by 10%.

9.31 Consider the BJT differential amplifier when fed with a common-mode voltage V_{CM} as shown in Fig. 9.15(a). As is often the case, the supply voltage V_{CC} may not be pure dc but might include a ripple component v_r of small amplitude and a frequency of 120 Hz (see Section 4.5). Thus the supply voltage becomes $V_{CC} + v_r$. Find the ripple component of the collector voltages, v_{C1} and v_{C2}, as well as of the difference output voltage $v_{od} \equiv v_{C2} - v_{C1}$. Comment on the differential amplifier response to this undesirable power-supply ripple.

D 9.32 Consider the differential amplifier of Fig. 9.14 and let the BJT β be very large:

(a) What is the largest input common-mode signal that can be applied while the BJTs remain comfortably in the active region with $v_{CB} = 0$?
(b) If the available power supply V_{CC} is 2.0 V, what value of IR_C should you choose in order to allow a common-mode input signal of ±1.0 V?
(c) For the value of IR_C found in (b), select values for I and R_C. Use the largest possible value for I subject to the

constraint that the base current of each transistor (when I divides equally) should not exceed 2 μA. Let $\beta = 100$.

9.33 To provide insight into the possibility of nonlinear distortion resulting from large differential input signals applied to the differential amplifier of Fig. 9.14, evaluate the normalized change in the current i_{E1}, $\Delta i_{E1}/I = \left(i_{E1} - (I/2)\right)/I$, for differential input signals v_{id} of 2, 5, 8, 10, 20, 30, and 40 mV. Provide a tabulation of the ratio $\left(\Delta i_{E1}/I\right)/v_{id}$, which represents the proportional transconductance gain of the differential pair, versus v_{id}. Comment on the linearity of the differential pair as an amplifier.

D 9.34 Design the circuit of Fig. 9.14 to provide a differential output voltage (i.e., one taken between the two collectors) of 1 V when the differential input signal is 10 mV. A current source of 1 mA and a positive supply of +5 V are available. What is the largest possible input common-mode voltage for which operation is as required? Assume $\alpha \simeq 1$.

***9.35** For the circuit in Fig. 9.14, assuming $\alpha = 1$ and $IR_C = 5$ V, use Eqs. (9.48) and (9.49) to find i_{C1} and i_{C2}, and hence determine $v_{od} = v_{C2} - v_{C1}$ for input differential signals $v_{id} \equiv v_{B1} - v_{B2}$ of 2 mV, 5 mV, 10 mV, 15 mV, 20 mV, 25 mV, 30 mV, 35 mV, and 40 mV. Plot v_{od} versus v_{id}, and hence comment on the amplifier linearity. As another way of visualizing linearity, determine the gain $\left(v_o/v_{id}\right)$ versus v_{id}. Comment on the resulting graph.

9.36 In a differential amplifier using a 1.5-mA emitter bias current source, the two BJTs are not matched. Rather, one has twice the emitter junction area of the other. For a differential input signal of zero volts, what do the collector currents become? What difference input is needed to equalize the collector currents? Assume $\alpha = 1$.

9.37 This problem explores the linearization of the transfer characteristics of the differential pair achieved by including emitter-degeneration resistances R_e in the emitters (see Fig. 9.17). Consider the case $I = 200$ μA with the transistors exhibiting $v_{BE} = 690$ mV at $i_C = 1$ mA and assume $\alpha \simeq 1$.

(a) With no emitter resistances R_e, what value of V_{BE} results when $v_{id} = 0$?
(b) With no emitter resistances R_e, use the large-signal model to find i_{C1} and i_{C2} when $v_{id} = 20$ mV.
(c) Now find the value of R_e that will result in the same i_{C1} and i_{C2} as in (b) but with $v_{id} = 200$ mV. Use the large-signal model.
(d) Calculate the effective transconductance G_m as the ratio of the difference current, $(i_{C1} - i_{C2})$, to v_{id} in the cases without and with the R_e's. By what factor is G_m reduced? How does this factor relate to the increase in v_{id}? Comment.

9.38 A BJT differential amplifier uses a 400-μA bias current. What is the value of g_m of each device? If β is 160, what is the differential input resistance?

D 9.39 Design the basic BJT differential amplifier circuit of Fig. 9.18 to provide a differential input resistance of at least 20 kΩ and a differential voltage gain of 100 V/V. The transistor β is specified to be at least 100. Specify I and R_C.

9.40 For a differential amplifier to which a total difference signal of 10 mV is applied, what is the equivalent signal to its corresponding CE half-circuit? If the emitter current source I is 200 μA, what is r_e of the half-circuit? For a load resistance of 10 kΩ in each collector, what is the half-circuit gain? What magnitude of signal output voltage would you expect at each collector? Between the two collectors?

9.41 A BJT differential amplifier is biased from a 0.5-mA constant-current source and includes a 400-Ω resistor in each emitter. The collectors are connected to V_{CC} via 10-kΩ resistors. A differential input signal of 0.1 V is applied between the two bases.

(a) Find the signal current in the emitters (i_e) and the signal voltage v_{be} for each BJT.
(b) What is the total emitter current in each BJT?
(c) What is the signal voltage at each collector? Assume $\alpha = 1$.
(d) What is the voltage gain realized when the output is taken between the two collectors?

D 9.42 Design a BJT differential amplifier to amplify a differential input signal of 0.1 V and provide a differential output signal of 2 V. To ensure adequate linearity, it is required to limit the signal amplitude across each base–emitter junction to a maximum of 5 mV. Another design requirement is that the differential input resistance be at least 100 kΩ. The BJTs available are specified to have $\beta \geq 100$. Give the circuit configuration and specify the values of all its components.

D 9.43 Design a bipolar differential amplifier such as that in Fig. 9.18 to operate from ±2.5 V power supplies and to provide differential gain of 60 V/V. The power dissipation in the quiescent state should not exceed 1 mW.

(a) Specify the values of I and R_C. What dc voltage appears at the collectors?
(b) If $\beta = 100$, what is the input differential resistance?
(c) For $v_{id} = 10$ mV, what is the signal voltage at each of the collectors?
(d) For the situation in (c), what is the maximum allowable value of the input common-mode voltage, V_{CM}? Recall that to maintain an *npn* BJT in saturation, v_B should not exceed v_C by more than 0.4 V.

D *9.44 In this problem we explore the trade-off between input common-mode range and differential gain in the design of the bipolar BJT. Consider the bipolar differential amplifier in Fig. 9.14 with the input voltages

$$v_{B1} = V_{CM} + (v_{id}/2)$$
$$v_{B2} = V_{CM} - (v_{id}/2)$$

(a) Bearing in mind that for a BJT to remain in the active mode, v_{BC} should not exceed 0.4 V, show that when v_{id} has a peak $\hat{v}_{id}$, the maximum input common-mode voltage $V_{CM\max}$ is given by

$$V_{CM\max} = V_{CC} + 0.4 - \frac{\hat{v}_{id}}{2} - A_d\left(V_T + \frac{\hat{v}_{id}}{2}\right)$$

(b) For the case $V_{CC} = 2.5$ V and $\hat{v}_{id} = 10$ mV, use the relationship above to determine $V_{CM\max}$ for the case $A_d =$ 50 V/V. Also find the peak output signal $\hat{v}_{od}$ and the required value of IR_C. Now if the power dissipation in the circuit is to be limited to 1 mW in the quiescent state (i.e., with $v_{id} = 0$), find I and R_C. (Remember to include the power drawn from the negative power supply $-V_{EE} = -2.5$ V.)
(c) If $V_{CM\max}$ is to be +1 V, and all other conditions remain the same, what maximum gain A_d is achievable?

9.45 For the differential amplifier of Fig. 9.14, let $V_{CC} =$ +5 V and $IR_C = 4$ V. Find the differential gain A_d. Sketch and clearly label the waveforms for the total collector voltages v_{C1} and v_{C2} and for $(v_{C2} - v_{C1})$ for the case:

$$v_{B1} = 1 + 0.005\ \sin(\omega t)$$
$$v_{B2} = 1 - 0.005\ \sin(\omega t)$$

9.46 Consider a bipolar differential amplifier in which the collector resistors R_C are replaced with simple current sources implemented using *pnp* transistors. Sketch the circuit and give its differential half-circuit. If $V_A = 20$ V for all transistors, find the differential voltage gain achieved.

9.47 For each of the emitter-degenerated differential amplifiers shown in Fig. P9.47, find the differential half-circuit and derive expressions for the differential gain A_d and differential input resistance R_{id}. For each circuit, what dc voltage appears across the bias current source(s) in the quiescent state (i.e., with $v_{id} = 0$)? Hence, which of the two circuits will allow a larger negative V_{CM}?

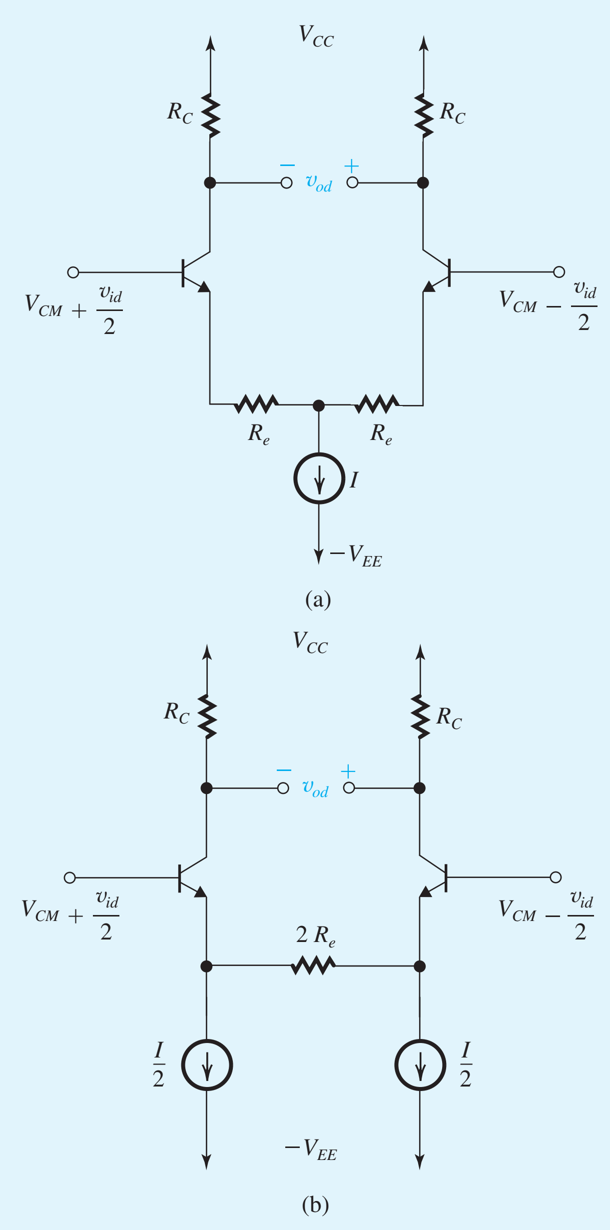


Figure P9.47

9.48 Consider a bipolar differential amplifier that, in addition to the collector resistances R_C, has a load resistance R_L connected between the two collectors. What does the differential gain A_d become?

9.49 A bipolar differential amplifier having resistance R_e inserted in series with each emitter (as in Fig. P9.47(a)) is biased with a constant current I. When both input terminals are grounded, the dc voltage measured across each R_e is found to be 4 V_T and that measured across each R_C is found to be 60 V_T. What differential voltage gain A_d do you expect the amplifier to have?

9.50 A bipolar differential amplifier with emitter-degeneration resistances R_e and R_e is fed with the arrangement shown in Fig. P9.50. Derive an expression for the overall differential voltage gain $G_v \equiv v_{od}/v_{sig}$. If R_{sig} is of such a value that $v_{id} = 0.5v_{sig}$, find the gain G_v in terms of R_C, r_e, R_e, and α. Now if β is doubled, by what factor does G_v increase?

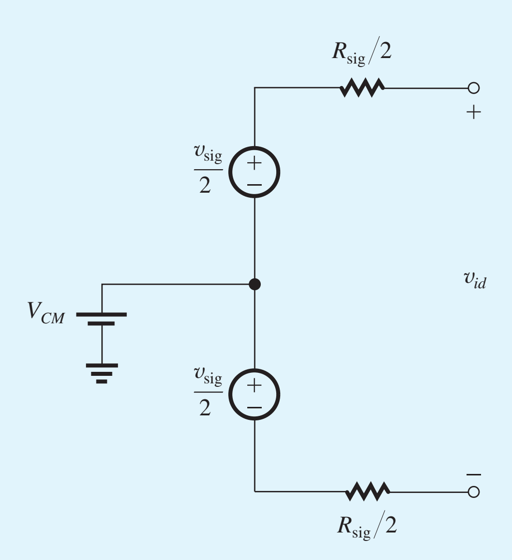


Figure P9.50

9.51 A particular differential amplifier operates from an emitter current source $I = 0.4$ mA. Each of the collector resistances $R_C = 20$ kΩ and a load resistance $R_L = 40$ kΩ is connected between the two collectors. If the amplifier is fed in the manner shown in Fig. P9.50 with $R_{sig} = 100$ kΩ, find the overall voltage gain. Let $\beta = 100$.

9.52 Find the voltage gain and the input resistance of the amplifier shown in Fig. P9.52 assuming $\beta = 100$.

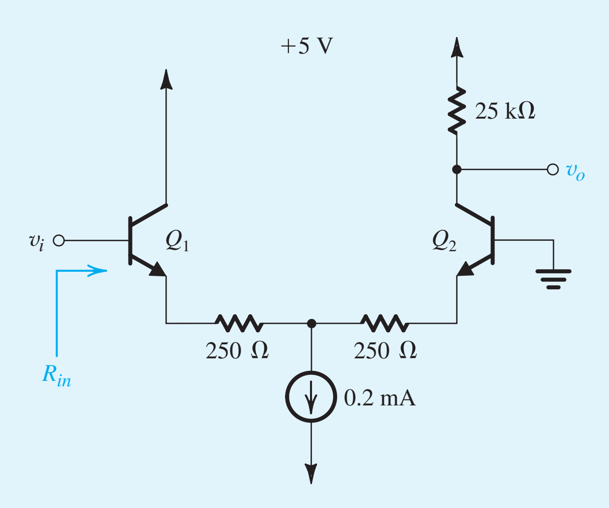


Figure P9.52

9.53 Find the voltage gain and input resistance of the amplifier in Fig. P9.53 assuming that $\beta = 100$.

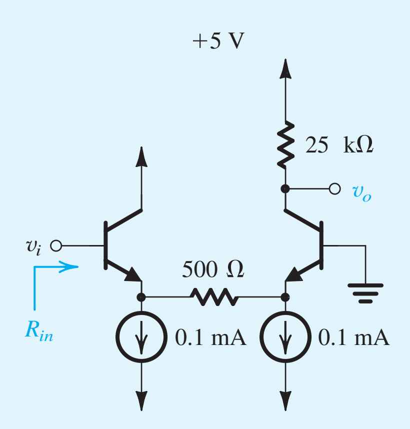


Figure P9.53

9.54 Derive an expression for the small-signal voltage gain v_o/v_i of the circuit shown in Fig. P9.54 in two different ways:

(a) as a differential amplifier
(b) as a cascade of a common-collector stage Q_1 and a common-base stage Q_2

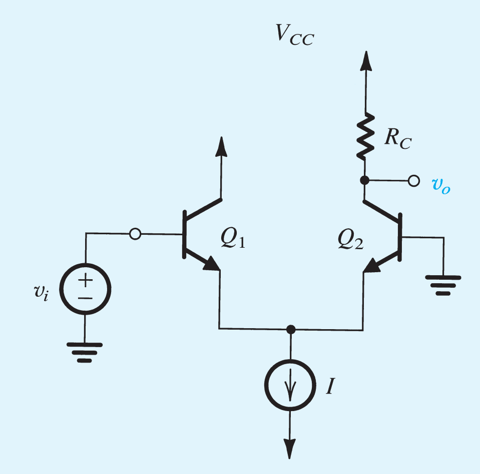


Figure P9.54

Assume that the BJTs are matched and have a current gain α, and neglect the Early effect. Verify that both approaches lead to the same result.

Section 9.3: Common-Mode Rejection

SIM **9.55** An NMOS differential pair is biased by a current source $I = 0.2$ mA having an output resistance $R_{SS} = 100$ kΩ. The amplifier has drain resistances $R_D = 10$ kΩ, using transistors with $k_n' W/L = 3$ mA/V^2, and r_o that is large. If the output is taken differentially and there is a 1% mismatch between the drain resistances, find $|A_d|$, $|A_{cm}|$, and CMRR.

9.56 For the differential amplifier shown in Fig. P9.2, let Q_1 and Q_2 have $k_p'(W/L) = 4$ mA/V^2, and assume that the bias current source has an output resistance of 30 kΩ. Find $|V_{OV}|$, g_m, $|A_d|$, $|A_{cm}|$, and the CMRR (in dB) obtained with the output taken differentially. The drain resistances are known to have a mismatch of 2%.

SIM **D *9.57** The differential amplifier in Fig. P9.57 utilizes a resistor R_{SS} to establish a 1-mA dc bias current. Note that this amplifier uses a single 5-V supply and thus the dc common-mode voltage V_{CM} cannot be zero. Transistors Q_1 and Q_2 have $k_n' W/L = 2.5$ mA/V^2, $V_t = 0.7$ V, and $\lambda = 0$.

(a) Find the required value of V_{CM}.
(b) Find the value of R_D that results in a differential gain A_d of 8 V/V.
(c) Determine the dc voltage at the drains.
(d) Determine the single-ended-output common-mode gain $\Delta V_{D1}/\Delta V_{CM}$. (*Hint:* You need to take $1/g_m$ into account.)
(e) Use the common-mode gain found in (d) to determine the change in V_{CM} that results in Q_1 and Q_2 entering the triode region.

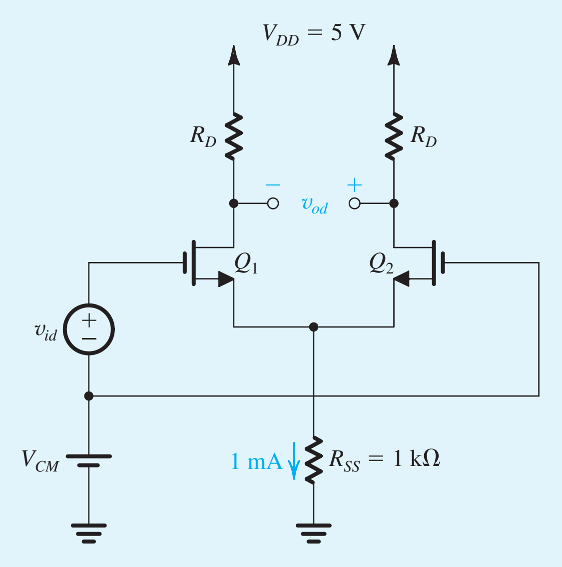


Figure P9.57

9.58 It can be shown that if the drain resistors of a MOS differential amplifier have a mismatch ΔR_D and if simultaneously the transconductances of Q_1 and Q_2 have a mismatch Δg_m, the common-mode gain is given by

$$A_{cm} \simeq \left(\frac{R_D}{2R_{SS}}\right)\left(\frac{\Delta g_m}{g_m} + \frac{\Delta R_D}{R_D}\right)$$

Note that this equation indicates that R_D can be deliberately varied to compensate for the initial variability in g_m and R_D, that is, to minimize A_{cm}.

In a MOS differential amplifier for which $R_D = 5$ kΩ and $R_{SS} = 25$ kΩ, the common-mode gain is measured and found to be 0.002 V/V. Find the percentage change required in one of the two drain resistors so as to reduce A_{cm} to zero (or close to zero).

D 9.59 It is required to design a MOS differential amplifier to have a CMRR of 80 dB. The only source of mismatch in the circuit is a 2% difference between the W/L ratios of the two transistors. Let $I = 100$ µA and assume that all transistors are operated at $V_{OV} = 0.2$ V. For the 0.18-µm CMOS fabrication process available, $V_A' = 5$ V/µm. What is the value of L required for the current-source transistor?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

D 9.60 A MOS differential amplifier utilizing a simple current source to provide the bias current I is found to have a CMRR of 60 dB. If it is required to raise the CMRR to 100 dB by adding a cascode transistor to the current source, what must the intrinsic gain A_0 of the cascode transistor be? If the cascode transistor is operated at $V_{OV} = 0.2$ V, what must its V_A be? If for the specific technology utilized $V_A' = 5$ V/μm, specify the channel length L of the cascode transistor.

9.61 The differential amplifier circuit of Fig. P9.61 utilizes a resistor connected to the negative power supply to establish the bias current I.

(a) For $v_{B1} = v_{id}/2$ and $v_{B2} = -v_{id}/2$, where v_{id} is a small signal with zero average, find the magnitude of the differential gain, $|v_o/v_{id}|$.
(b) For $v_{B1} = v_{B2} = v_{icm}$, where v_{icm} has a zero average, find the magnitude of the common-mode gain, $|v_o/v_{icm}|$.
(c) Calculate the CMRR.
(d) If $v_{B1} = 0.1 \sin 2\pi \times 60t + 0.005 \sin 2\pi \times 1000t$, volts, and $v_{B2} = 0.1 \sin 2\pi \times 60t - 0.005 \sin 2\pi \times 1000t$, volts, find v_o.

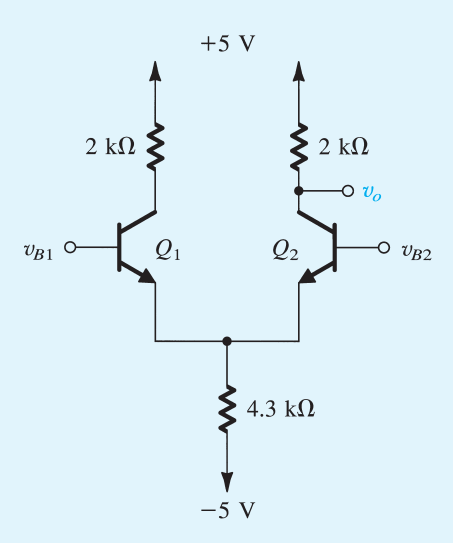


Figure P9.61

9.62 For the differential amplifier shown in Fig. P9.62, identify and sketch the differential half-circuit and the common-mode half-circuit. Find the differential gain, the differential input resistance, the common-mode gain assuming the resistances R_C have 1% tolerance, and the common-mode input resistance. For these transistors, $\beta = 100$ and $V_A = 100$ V.

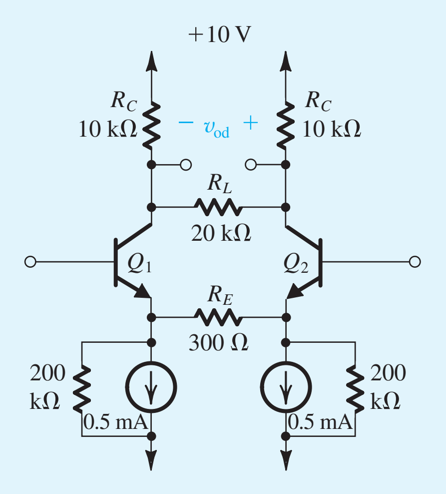


Figure P9.62

9.63 Consider the basic differential circuit in which the transistors have $\beta = 100$ and $V_A = 100$ V, with $I = 0.2$ mA, $R_{EE} = 500$ kΩ, and $R_C = 25$ kΩ. The collector resistances are matched to within 1%. Find:

(a) the differential gain
(b) the differential input resistance
(c) the common-mode gain
(d) the common-mode rejection ratio
(e) the common-mode input resistance

9.64 In a bipolar differential-amplifier circuit, the bias current generator consists of a simple common-emitter transistor operating at 200 μA. For this transistor, and those used in the differential pair, $V_A = 20$ V and $\beta = 50$. What common-mode input resistance would result? Assume $R_C \ll r_o$.

9.65 A bipolar differential amplifier with $I = 0.5$ mA utilizes transistors for which $V_A = 50$ V and $\beta = 100$. The collector resistances $R_C = 5$ kΩ and are matched to within 10%. Find:

(a) the differential gain
(b) the common-mode gain and the CMRR if the bias current I is generated using a simple current mirror
(c) the common-mode gain and the CMRR if the bias current I is generated using a Wilson mirror. (Refer to Eq. 8.95 for R_o of the Wilson mirror.)

D 9.66 It is required to design a differential amplifier to provide the largest possible signal to a pair of 10-kΩ load

resistances. The input differential signal is a sinusoid of 5-mV peak amplitude, which is applied to one input terminal while the other input terminal is grounded. The power supply V_{CC} available is 5 V. To determine the required bias current I, derive an expression for the total voltage at each of the collectors in terms of V_{CC} and I in the presence of the input signal. Then impose the condition that both transistors remain well out of saturation with a minimum v_{CB} of approximately 0 V. Thus determine the required value of I. For this design, what differential gain is achieved? What is the amplitude of the signal voltage obtained between the two collectors? Assume $\alpha \simeq 1$.

D *9.67 Design a BJT differential amplifier that provides two single-ended outputs (at the collectors). The amplifier is to have a differential gain (to each of the two outputs) of at least 100 V/V, a differential input resistance $\geq$ 10 kΩ, and a common-mode gain (to each of the two outputs) no greater than 0.1 V/V. Use a 2-mA current source for biasing. Give the complete circuit with component values and suitable power supplies that allow for ±2 V swing at each collector. Specify the minimum value that the output resistance of the bias current source must have. If the current source is realized by a simple mirror, what must the minimum value of V_A be? The BJTs available have $\beta \geq 100$. What is the value of the input common-mode resistance when the bias source has the lowest acceptable output resistance?

9.68 When the output of a BJT differential amplifier is taken differentially, its CMRR is found to be 34 dB higher than when the output is taken single-endedly. If the only source of common-mode gain when the output is taken differentially is the mismatch in collector resistances, what must this mismatch be (in percent)?

***9.69** In a particular BJT differential amplifier, a production error results in one of the transistors having an emitter– base junction area that is twice that of the other. With the inputs grounded, how will the emitter bias current split between the two transistors? If the output resistance of the current source is 500 kΩ and the resistance in each collector (R_C) is 12 kΩ, find the common-mode gain obtained when the output is taken differentially. Assume $\alpha \simeq 1$. [*Hint*: The CM signal current v_{icm}/R_{EE} will split between Q_1 and Q_2 in the same ratio as the bias current I does.]

Section 9.4: DC Offset

D 9.70 An NMOS differential pair is to be used in an amplifier whose drain resistors are 10 kΩ ± 1%. For the pair, $k_n' W/L = 4$ mA/V^2. A decision is to be made concerning the bias current I to be used, whether 160 μA or 360 μA. Contrast the differential gain and input offset voltage for the two possibilities.

D 9.71 An NMOS differential amplifier for which the MOSFETs have a transconductance parameter k_n and whose drain resistances R_D have a mismatch ΔR_D is biased with a current I.

(a) Find expressions for A_d and V_{OS} in terms of k_n, R_D, $\Delta R_D/R_D$, and I. Use these expressions to relate V_{OS} and A_d.
(b) If $k_n = 4$ mA/V^2, $R_D = 10$ kΩ, and $\Delta R_D/R_D = 0.02$, find the maximum gain realized if V_{OS} is to be limited to 1 mV, 2 mV, 3 mV, 4 mV, and 5 mV. For each case, give the value of the required bias current I. Note the trade-off between gain and offset voltage.

D 9.72 An NMOS amplifier, whose designed operating point is at $V_{OV} = 0.3$ V, is suspected to have a variability of V_t of ±5 mV, and of W/L and R_D (independently) of ±1%. What is the worst-case input offset voltage you would expect to find? What is the major contribution to this total offset? If you used a variation of one of the drain resistors to reduce the output offset to zero and thereby compensate for the uncertainties (including that of the other R_D), what percentage change from nominal would you require?

9.73 An NMOS differential pair operating at a bias current I of 100 μA uses transistors for which $k_n' = 200$ μA/V^2 and $W/L = 10$. Find the three components of input offset voltage under the conditions that $\Delta R_D/R_D = 4\%$, $\Delta(W/L)/(W/L) = 4\%$, and $\Delta V_t = 5$ mV. In the worst case, what might the total offset be? For the usual case of the three effects being independent, what is the offset likely to be?

9.74 A bipolar differential amplifier uses two well-matched transistors, but collector load resistors that are mismatched by 10%. What input offset voltage is required to reduce the differential output voltage to zero?

9.75 A bipolar differential amplifier uses two transistors whose scale currents I_S differ by 10%. If the two collector resistors are well matched, find the resulting input offset voltage.

9.76 Modify Eq. (9.114) for the case of a differential amplifier having a resistance R_E connected in the emitter of each transistor. Let the bias current source be I.

9.77 A differential amplifier uses two transistors whose β values are β_1 and β_2. If everything else is matched, show that

the input offset voltage is approximately $V_T[(1/\beta_1) - (1/\beta_2)]$. Evaluate V_{OS} for $\beta_1 = 50$ and $\beta_2 = 100$.

9.78 Two possible differential amplifier designs are considered, one using BJTs and the other MOSFETs. In both cases, the collector (drain) resistors are maintained within ±2% of nominal value. The MOSFETs are operated at $V_{OV} = 200$ mV. What input offset voltage results in each case? What does the MOS V_{OS} become if the devices are increased in width by a factor of 4 while the bias current is kept constant?

***9.79** A differential amplifier uses two transistors having V_A values of 100 V and 200 V. If everything else is matched, find the resulting input offset voltage. Assume that the two transistors are intended to be biased at a V_{CE} of about 10 V.

***9.80** A differential amplifier is fed in a balanced or push–pull manner, and the source resistance in series with each base is R_s. Show that a mismatch ΔR_s between the values of the two source resistances gives rise to an input offset voltage of approximately $(I/2\beta)\Delta R_s / [1 + (g_m R_s)/\beta]$.

9.81 One approach to "offset correction" involves the adjustment of the values of R_{C1} and R_{C2} so as to reduce the differential output voltage to zero when both input terminals are grounded. This offset-nulling process can be accomplished by utilizing a potentiometer in the collector circuit, as shown in Fig. P9.81. We wish to find the potentiometer setting, represented by the fraction x of its value connected in series with R_{C1}, that is required for nulling the output offset voltage that results from:

(a) R_{C1} being 4% higher than nominal and R_{C2} 4% lower than nominal

(b) Q_1 having an area 5% larger than nominal, while Q_2 has area 5% smaller than nominal.

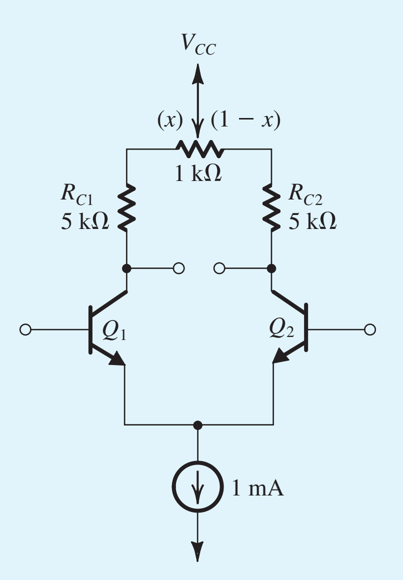


Figure P9.81

9.82 A differential amplifier for which the total emitter bias current is 400 μA uses transistors for which β is specified to lie between 80 and 200. What is the largest possible input bias current? The smallest possible input bias current? The largest possible input offset current?

****9.83** In a particular BJT differential amplifier, a production error results in one of the transistors having an emitter–base junction area twice that of the other. With both inputs grounded, find the current in each of the two transistors and hence the dc offset voltage at the output, assuming that the collector resistances are equal. Use small-signal analysis to find the input voltage that would restore current balance to the differential pair. Repeat using large-signal analysis and compare results.

D 9.84 A large fraction of mass-produced differential-amplifier modules employing 20-kΩ collector resistors is found to have an input offset voltage ranging from +2 mV to –2 mV. By what amount must one collector resistor be adjusted to reduce the input offset to zero? If an adjustment mechanism is devised that raises one collector resistance while correspondingly lowering the other, what resistance change is needed? If a potentiometer connected as shown in Fig. P9.81 is used, what value of potentiometer resistance (specified to 1 significant digit) is needed? Assume that the offset is entirely due to the finite tolerance of R_C.

Section 9.5: The Differential Amplifier with a Current-Mirror Load

9.85 The differential amplifier of Fig. 9.32(a) is measured and found to have a short-circuit transconductance of 2 mA/V. A differential input signal is applied and the output voltage is measured with a load resistance R_L connected. It is found that when R_L is reduced from ∞ to 20 kΩ, the magnitude of the output signal is reduced by half. What do you estimate R_o and A_d (with R_L disconnected) to be?

9.86 A current-mirror-loaded NMOS differential amplifier is fabricated in a technology for which $|V_A'| = 5$ V/μm. All the

transistors have $L = 0.5\ \mu$m. If the differential-pair transistors are operated at $V_{OV} = 0.25$ V, what open-circuit differential gain is realized?

9.87 The differential amplifier of Fig. 9.32(a) is biased with $I = 200\ \mu$A. All transistors have $L = 0.5\ \mu$m, and Q_1 and Q_2 have $W/L = 50$. The circuit is fabricated in a process for which $\mu_n C_{ox} = 200\ \mu\text{A/V}^2$ and $|V_A'| = 5$ V/μm. Find $g_{m1,2}$, r_{o2}, r_{o4}, and A_d.

D 9.88 In a current-mirror-loaded differential amplifier of the form shown in Fig. 9.32(a), all transistors are characterized by $k'W/L = 4\ \text{mA/V}^2$, and $|V_A| = 5$ V. Find the bias current I for which the gain $v_o/v_{id} = 20$ V/V.

D 9.89 Consider a current-mirror-loaded differential amplifier such as that shown in Fig. 9.32(a) with the bias current source implemented with the modified Wilson mirror of Fig. P9.89 with $I = 200\ \mu$A. The transistors have $|V_t| = 0.5$V and $k'W/L = 5\ \text{mA/V}^2$. What is the lowest value of the total power supply $(V_{DD} + V_{SS})$ that allows each transistor to operate with $|V_{DS}| \geq |V_{GS}|$?

***9.90** (a) Sketch the circuit of a current-mirror-loaded MOS differential amplifier in which the input transistors are cascoded and a cascode current mirror is used for the load.

(b) Show that if all transistors are operated at an overdrive voltage V_{OV} and have equal Early voltages $|V_A|$, the gain is given by

$$A_d = 2(V_A/V_{OV})^2$$

Evaluate the gain for $V_{OV} = 0.20$ V and $V_A = 10$ V.

9.91 Figure P9.91 shows the current-mirror-loaded MOS differential amplifier prepared for small-signal analysis. We have "pulled out" r_o of each transistor; thus, the current in the drain of each transistor will be $g_m v_{gs}$. To help the reader, we have already indicated approximate values for some of the node voltages. For instance, the output voltage $v_o = \frac{1}{2}(g_m r_o) v_{id}$, which we have derived in the text. The voltage at the common sources has been found to be approximately $+v_{id}/4$, which is very far from the virtual ground one might assume. Also, the voltage at the gate of the mirror is approximately $-v_{id}/4$, confirming our contention that the voltage there is vastly different from the output voltage, hence the lack of balance in the circuit and the unavailability of a differential half-circuit. Find the currents labeled i_1 to i_{13} in terms of $(g_m v_{id})$. Determine their values in the sequence of their numbering and assume $g_m r_o \gg 1$. Note that all transistors are assumed to be operating at the same $|V_{OV}|$. Write the current values on the circuit diagram and reflect on the results.

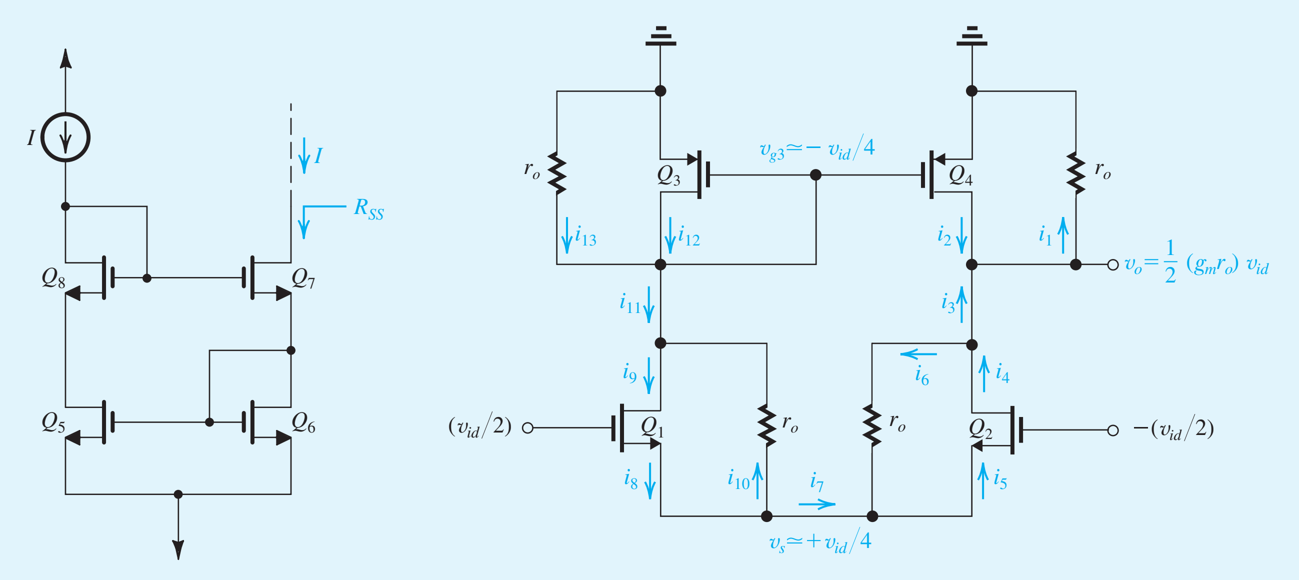


Figure P9.89 Figure P9.91

9.92 A current-mirror-loaded NMOS differential amplifier operates with a bias current I of 200 μA. The NMOS transistors are operated at $V_{OV} = 0.2$ V and the PMOS devices at $|V_{OV}| = 0.3$ V. The Early voltages are 20 V for the NMOS and 12 V for the PMOS transistors. Find G_m, R_o, and A_d. For what value of load resistance is the gain reduced by a factor of 2?

9.93 This problem investigates the effect of transistor mismatches on the input offset voltage of the current-mirror-loaded MOS differential amplifier of Fig. 9.32(a). For this purpose, ground both input terminals and short-circuit the output node to ground.
(a) If the amplifying transistors Q_1 and Q_2 exhibit a W/L mismatch of $\Delta(W/L)_A$, find the resulting short-circuit output current and hence show that the corresponding V_{OS} is given by

$$V_{OS1} = (V_{OV}/2)\frac{\Delta(W/L)_A}{(W/L)_A}$$

where V_{OV} is the overdrive voltage at which Q_1 and Q_2 are operating.
(b) Repeat for a mismatch $\Delta(W/L)_M$ in the W/L ratios of the mirror transistor Q_3 and Q_4 to show that the corresponding V_{OS} is given by

$$V_{OS2} = (V_{OV}/2)\frac{\Delta(W/L)_M}{(W/L)_M}$$

where V_{OV} is the overdrive voltage at which Q_1 and Q_2 are operating.
(c) For a circuit in which all transistors are operated at $|V_{OV}| = 0.2$ V and all W/L ratios are accurate to within ±1% of nominal, find the worst-case total offset voltage V_{OS}.

9.94 The differential amplifier in Fig. 9.36(a) is operated with $I = 500$ μA, with devices for which $V_A = 10$ V and $\beta = 100$. What differential input resistance, output resistance, short-circuit transconductance, and open-circuit voltage gain would you expect? What will the voltage gain be if the input resistance of the subsequent stage is equal to R_{id} of this stage?

9.95 A bipolar differential amplifier having a simple *pnp* current-mirror load is found to have an input offset voltage of 2 mV. If the offset is attributable entirely to the finite β of the *pnp* transistors, what must β_P be?

9.96 For the current-mirror-loaded bipolar differential pair, replacing the simple current-mirror load by the base-current-compensated mirror of Fig. 8.11, find the expected systematic input offset voltage. Evaluate V_{OS} for $\beta_P = 50$.

9.97 For the current-mirror-loaded bipolar differential pair, replacing the simple current-mirror load by the Wilson mirror of Fig. 8.40(a), find the expected systematic input offset voltage. Evaluate V_{OS} for $\beta_P, = 50$.

9.98 Figure P9.98 shows a differential cascode amplifier with an active load formed by a Wilson current mirror. Utilizing the expressions derived in Chapter 8 for the output resistance of a bipolar cascode and the output resistance of the Wilson mirror, and assuming all transistors to be identical, show that the differential voltage gain A_d is given approximately by

$$A_d = \frac{1}{3}\beta g_m r_o$$

Evaluate A_d for the case of $\beta = 100$ and $V_A = 20$ V.

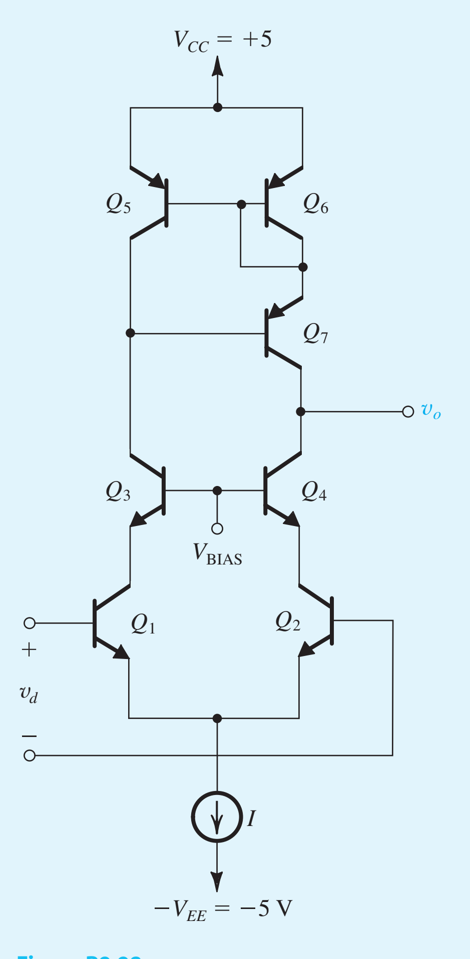


Figure P9.98

D 9.99 Consider the bias design of the Wilson-loaded cascode differential amplifier shown in Fig. P9.98.

(a) What is the largest signal voltage possible at the output without Q_7 saturating? Assume that the CB junction conducts when the voltage across it exceeds 0.4 V.
(b) What should the dc bias voltage established at the output (by an arrangement not shown) be in order to allow for positive output signal swing of 1.5 V?
(c) What should the value of V_{BIAS} be in order to allow for a negative output signal swing of 1.5 V?
(d) What is the upper limit on the input common-mode voltage v_{CM}?

****9.100** Figure P9.100 shows a modified cascode differential amplifier. Here Q_3 and Q_4 are the cascode transistors. However, the manner in which Q_3 is connected with its base current feeding the current mirror Q_7–Q_8 results in very interesting input properties. Note that for simplicity the circuit is shown with the base of Q_2 grounded. Assume that all transistors have equal β's.

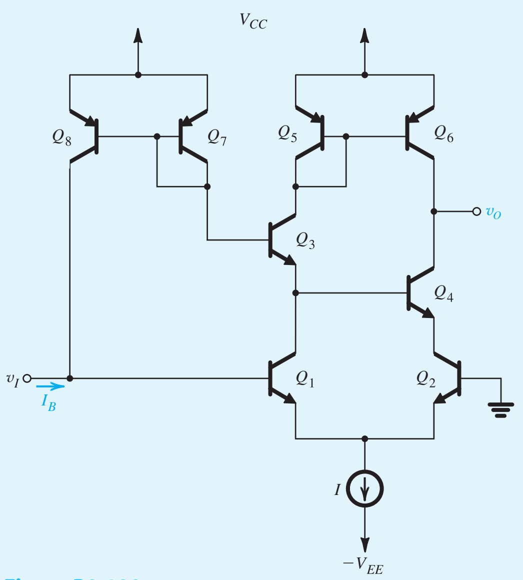


Figure P9.100

(a) With $v_I = 0$ V dc, find the input bias current I_B assuming all transistors have equal value of β. Compare with the case without the Q_7–Q_8 connection.
(b) With $v_I = 0$ V (dc) + v_{id}, find the input signal current i_i and hence the input differential resistance R_{id}. Compare with the case without the Q_7–Q_8 connection. By what factor does R_{id} increase?

9.101 For the folded-cascode differential amplifier of Fig. 9.38, find the value of V_{BIAS} that results in the largest possible positive output swing, while keeping Q_3, Q_4, and the *pnp* transistors that realize the current sources out of saturation. Assume $V_{CC} = V_{EE} = 5$ V. If the dc level at the output is 0 V, find the maximum allowable output signal swing. For $I = 0.5$ mA, $\beta_P = 50$, $\beta_N = 100$, and $V_A = 100$ V find G_m, R_{o4}, R_{o5}, R_o, and A_d.

9.102 For the BiCMOS differential amplifier in Fig. P9.102 let $V_{DD} = V_{SS} = 3$ V, $I = 0.2$ mA, $k_p' W/L = 6.4$ mA/V^2; $|V_A|$ for *p*-channel MOSFETs is 10 V, $|V_A|$ for *npn* transistors is 30 V. Find G_m, R_o, and A_d.

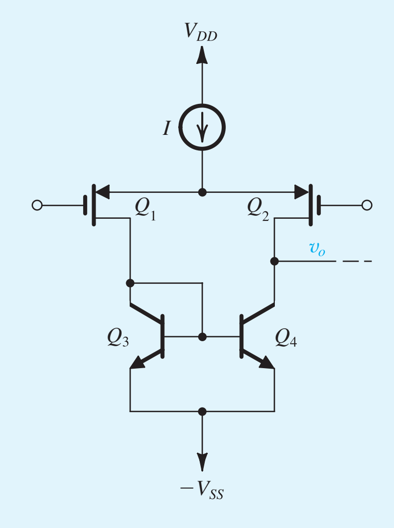


Figure P9.102

SIM D 9.103 It is required to design the current-mirror-loaded differential MOS amplifier of Fig. 9.32 to obtain a differential gain of 50 V/V. The technology available provides $\mu_n C_{ox} = 4\mu_p C_{ox} = 400\ \mu A/V^2$, $|V_t| = 0.5$ V, and $|V_A'| = 20$ V/μm and operates from ±1 V supplies. Use a bias current $I = 200$ μA and operate all devices at $|V_{OV}| = 0.2$ V.

(a) Find the W/L ratios of the four transistors.
(b) Specify the channel length required of all transistors.
(c) If $V_{ICM} = 0$, what is the allowable range of v_O?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

(d) If I is delivered by a simple NMOS current source operated at the same V_{OV} and having the same channel length as the other four transistors, determine the CMRR obtained.

9.104 Consider the current-mirror-loaded MOS differential amplifier of Fig. 9.32(a) in two cases:

(a) Current source I is implemented with a simple current mirror.

(b) Current source I is implemented with the modified Wilson current mirror shown in Fig. P9.104.

Recalling that for the simple mirror $R_{SS} = r_o|_{Q_S}$ and for the Wilson mirror $R_{SS} \simeq g_{m7}r_{o7}r_{o5}$, and assuming that all transistors have the same $|V_A|$ and $k'W/L$, show that for case (a)

$$\text{CMRR} = 2\left(\frac{V_A}{V_{OV}}\right)^2$$

and for case (b)

$$\text{CMRR} = 2\sqrt{2}\left(\frac{V_A}{V_{OV}}\right)^3$$

where V_{OV} is the overdrive voltage that corresponds to a drain current of $I/2$. For $k'W/L = 4\text{ mA/V}^2$, $I = 160\ \mu\text{A}$, and $|V_A| = 5$ V, find CMRR for both cases.

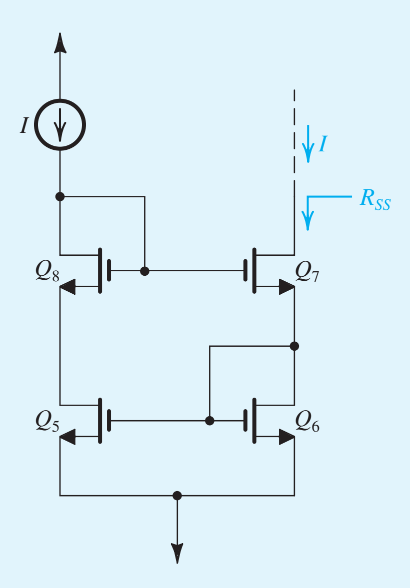


Figure P9.104

9.105 The MOS differential amplifier of Fig. 9.32(a) is biased with a simple current mirror delivering $I = 200\ \mu\text{A}$. All transistors are operated at $V_{OV} = 0.2$ V and have $V_A = 5$ V. Find G_m, R_o, A_d, R_{SS}, G_{mcm}, R_{im}, A_m, R_{om}, R_{o2}, A_{cm}, and CMRR.

9.106 A current-mirror-loaded MOS differential amplifier is found to have a differential voltage gain A_d of 30 V/V. Its bias current source has an output resistance $R_{SS} = 45$ kΩ. The current mirror utilized has a current gain A_m of 0.98 A/A and an output resistance R_{om} of 45 kΩ. If the common-mode output resistances of the amplifier, R_{o1} and R_{o2}, are very large, find A_{cm} and CMRR.

9.107 A current-mirror-loaded MOS differential amplifier is found to have a differential voltage gain A_d of 50 V/V and a CMRR of 60 dB. If the output resistance of the bias current source is 20 kΩ and the output resistance of the current-mirror load is 20 kΩ, what is the expected magnitude of the deviation from unity of the current gain of the load mirror?

D *9.108 Design the circuit of Fig. 9.36(a) using a basic current mirror to implement the current source I. It is required that the short-circuit transconductance be 5 mA/V. Use ±5-V power supplies and BJTs that have $\beta = 100$ and $V_A = 100$ V. Give the complete circuit with component values and specify the differential input resistance R_{id}, the output resistance R_o, the open-circuit voltage gain A_d, the input bias current, the input common-mode range, the common-mode gain, and the CMRR.

D *9.109 Repeat the design of the amplifier specified in Problem 9.108 utilizing a Widlar current source (Fig. 8.42) to supply the bias current. Assume that the largest resistance available is 2 kΩ.

9.110 A bipolar differential amplifier such as that shown in Fig. 9.36(a) has $I = 0.4$ mA, $V_A = 40$ V, and $\beta = 150$. Find G_m, R_o, A_d, and R_{id}. If the bias current source is implemented with a simple *npn* current mirror, find R_{EE}, A_{cm}, and CMRR. If the amplifier is fed differentially with a source having a total of 30 kΩ resistance (i.e., 15 kΩ in series with the base lead of each of Q_1 and Q_2), find the overall differential voltage gain.

9.111 For the current-mirror-loaded differential pair in Fig. P9.111, find:

(a) differential input resistance, R_{id}
(b) A_d
(c) CMRR

Assume $\beta = 100, |V_{BE}| = 0.7$ V, and $|V_A| = 60$ V.

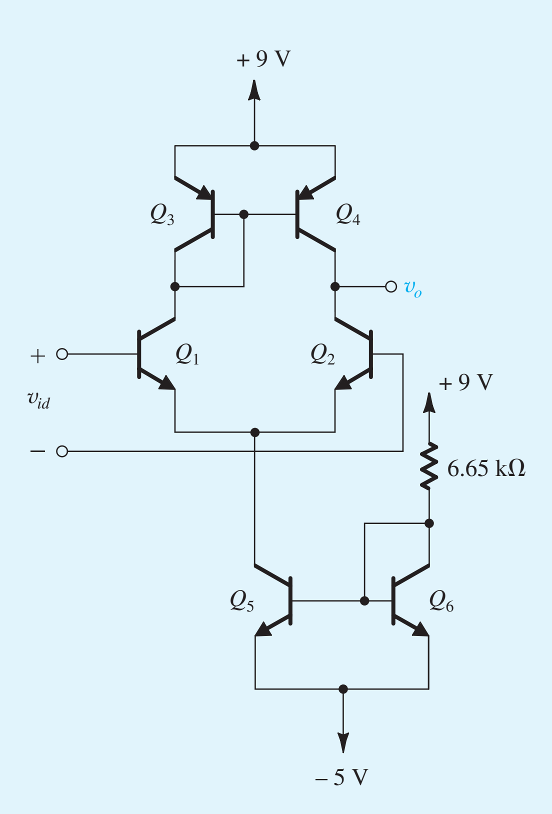


Figure P9.111

9.112 For the current-mirror-loaded differential amplifier in Fig. P9.112, find:

(a) differential input resistance, R_{id}
(b) A_d
(c) CMRR

Assume $\beta = 100, |V_{BE}| = 0.7$ V, $|V_A| = 60$ V, $V_t = 0.7$ V, and $k'(W/L) = 2$ mA/V^2.

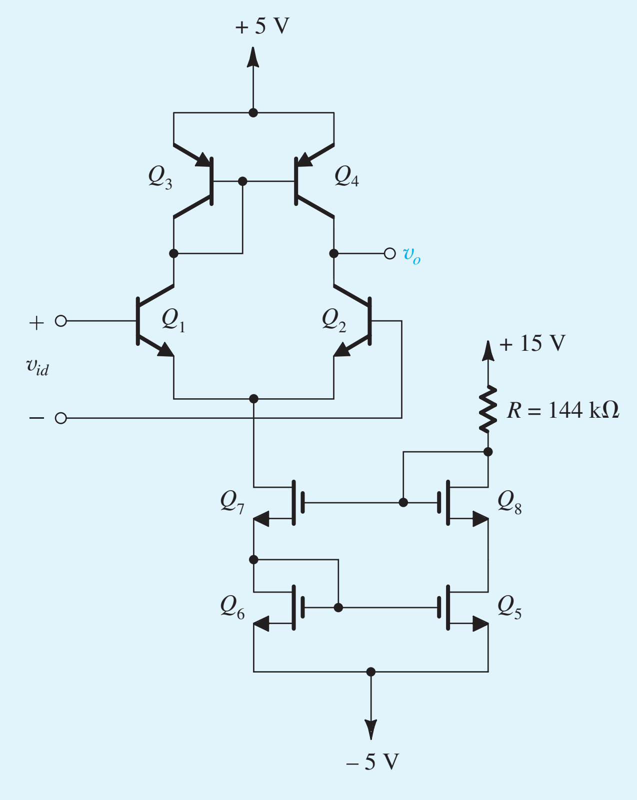


Figure P9.112

Section 9.6: Multistage Amplifiers

9.113 Consider the circuit in Fig. 9.40 with the device geometries (in μm) shown in Table P9.113. Let $I_{REF} = 225\ \mu$A, $|V_t| = 0.75$ V for all devices, $\mu_n C_{ox} = 180\ \mu$A/V^2, $\mu_p C_{ox} = 60\ \mu$A/V^2, $|V_A| = 9$ V for all devices, $V_{DD} = V_{SS} = 1.5$ V. Determine the width of Q_6, W, that will ensure that the op amp will not have a systematic offset voltage. Then, for all devices evaluate I_D, $|V_{OV}|$, $|V_{GS}|$, g_m, and r_o. Provide your results in a table similar to Table 9.1. Also find A_1, A_2, the open-loop voltage gain, the input

Table P9.113

Transistor	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8
W/L	30/0.5	30/0.5	10/0.5	10/0.5	60/0.5	W/0.5	60/0.5	60/0.5

common-mode range, and the output voltage range. Neglect the effect of V_A on the bias currents.

(d) With $v_A = v_{id}/2$ and $v_B = -v_{id}/2$, find the voltage gain v_o/v_{id}. Assume an Early voltage of 6 V.

D 9.114 The two-stage CMOS op amp in Fig. P9.114 is fabricated in a 0.18-μm technology having $k_n' = 4k_p' = 400\ \mu\text{A/V}^2$, $V_{tn} = -V_{tp} = 0.4$ V.

(a) With A and B grounded, perform a dc design that will result in each of Q_1, Q_2, Q_3, and Q_4 conducting a drain current of 100 μA and each of Q_6 and Q_7 a current of 200 μA. Design so that all transistors operate at 0.2-V overdrive voltages. Specify the W/L ratio required for each MOSFET. Present your results in tabular form. What is the dc voltage at the output (ideally)?
(b) Find the input common-mode range.
(c) Find the allowable range of the output voltage.

D *9.115 In a particular design of the CMOS op amp of Fig. 9.40 the designer wishes to investigate the effects of increasing the W/L ratio of both Q_1 and Q_2 by a factor of 4. Assuming that all other parameters are kept unchanged, refer to Example 9.6 to help you answer the following questions:

(a) Find the resulting change in $|V_{OV}|$ and in g_m of Q_1 and Q_2.
(b) What change results in the voltage gain of the input stage? In the overall voltage gain?
(c) What is the effect on the input offset voltages? (You might wish to refer to Section 9.4).

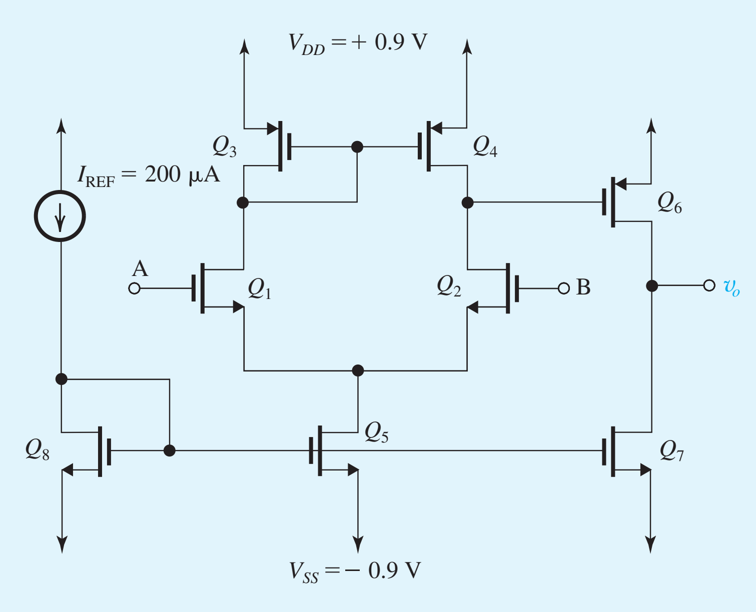


Figure P9.114

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

9.116 Consider the amplifier of Fig. 9.40, whose parameters are specified in Example 9.6. If a manufacturing error results in the *W*/*L* ratio of Q_7 being 48/0.8, find the current that Q_7 will now conduct. Thus find the systematic offset voltage that will appear at the output. (Use the results of Example 9.6.) Assuming that the open-loop gain will remain approximately unchanged from the value found in Example 9.6, find the corresponding value of input offset voltage, V_{OS}.

9.117 Consider the input stage of the CMOS op amp in Fig. 9.40 with both inputs grounded. Assume that the two sides of the input stage are perfectly matched except that the threshold voltages of Q_3 and Q_4 have a mismatch ΔV_t. Show that a current $g_{m3}\Delta V_t$ appears at the output of the first stage. What is the corresponding input offset voltage?

9.118 The two-stage op amp in Figure P9.114 is fabricated in a 65-nm technology having $k_n' = 5.4 \times k_p' = 540\ \mu\text{A/V}^2$ and $V_{tn} = -V_{tp} = 0.35$ V. The amplifier is operated with $V_{DD} = +1.2$ V and $V_{SS} = 0$ V.

(a) With A and B at a dc voltage of $V_{DD}/2$, perform a dc design that will result in each of Q_1, Q_2, Q_3, and Q_4 conducting a drain current of 200 μA and each of Q_6 and Q_7 conducting a current of 400 μA. Design so that all transistors operate at 0.15-V overdrive voltages. Specify the *W*/*L* ratio required for each MOSFET. Present all results in a table.

(b) Find the input common-mode range.

(c) Find the allowable range of the output voltage.

(d) With $v_A = v_{id}/2$ and $v_B = -v_{id}/2$, find the voltage gain v_o/v_{id}. Assume an Early voltage of 1.8 V.

***9.119** Figure P9.119 shows a bipolar op-amp circuit that resembles the CMOS op amp of Fig. 9.40. Here, the input differential pair Q_1–Q_2 is loaded in a current mirror formed by Q_3 and Q_4. The second stage is formed by the current-source-loaded common-emitter transistor Q_5. Unlike the CMOS circuit, here there is an output stage formed by the emitter follower Q_6. The function of capacitor C_C will be explained later, in Chapter 11. All transistors have $\beta = 100$, $|V_{BE}| = 0.7$ V, and $r_o = \infty$.

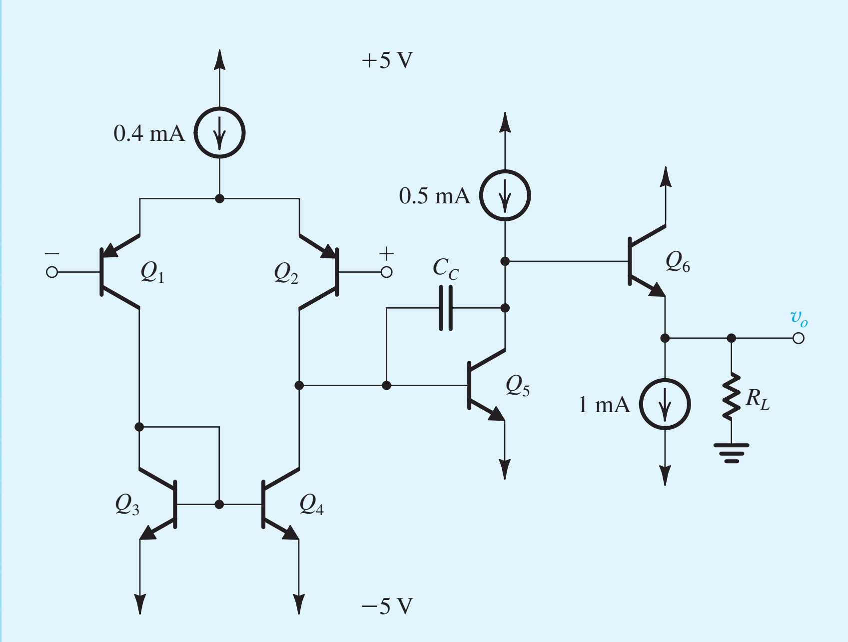


Figure P9.119

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

(a) For inputs grounded and output held at 0 V (by negative feedback, not shown) find the emitter currents of all transistors.
(b) Calculate the gain of the amplifier with $R_L = 1$ kΩ.

9.120 A BJT differential amplifier, biased to have $r_e = 50\ \Omega$ and utilizing two 50-Ω emitter resistors and 5-kΩ loads, drives a second differential stage biased to have $r_e = 25\ \Omega$. All BJTs have $\beta = 100$. What is the voltage gain of the first stage? Also find the input resistance of the first stage, and the current gain from the input of the first stage to the collectors of the second stage.

9.121 In the multistage amplifier of Fig. 9.41, emitter resistors are to be introduced—100 Ω in the emitter lead of each of the first-stage transistors and 25 Ω for each of the second-stage transistors. What is the effect on input resistance, the voltage gain of the first stage, and the overall voltage gain? Use the bias values found in Example 9.7.

D 9.122 Consider the circuit of Fig. 9.41 and its output resistance. Which resistor has the most effect on the output resistance? What should this resistor be changed to if the output resistance is to be reduced by a factor of 2? What will the amplifier gain become after this change? What other change can you make to restore the amplifier gain to approximately its prior value?

D 9.123 (a) If, in the multistage amplifier of Fig. 9.41, the resistor R_5 is replaced by a constant-current source $\simeq 1$ mA, such that the bias situation is essentially unaffected, what does the overall voltage gain of the amplifier become? Assume that the output resistance of the current source is very high. Use the results of Example 9.8.
(b) With the modification suggested in (a), what is the effect of the change on output resistance? What is the overall gain of the amplifier when loaded by 100 Ω to ground? The original amplifier (before modification) has an output resistance of 152 Ω and a voltage gain of 8513 V/V. What is its gain when loaded by 100 Ω? Comment. Use $\beta = 100$.

***9.124** Figure P9.124 shows a three-stage amplifier in which the stages are directly coupled. The amplifier, however, utilizes bypass capacitors, and, as such, its frequency response falls off at low frequencies. For our purposes here, we shall assume that the capacitors are large enough to act as perfect short circuits at all signal frequencies of interest.

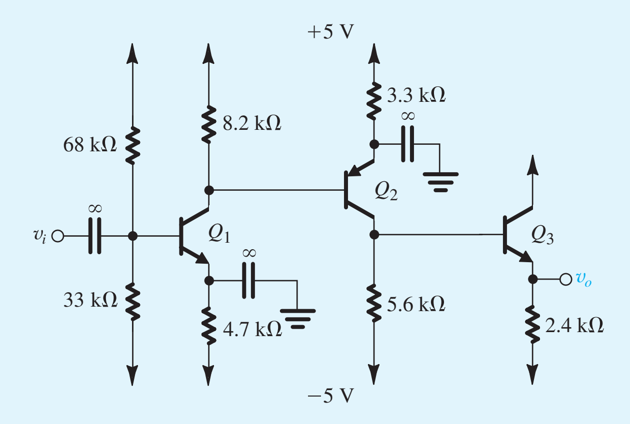


Figure P9.124

(a) Find the dc bias current in each of the three transistors. Also find the dc voltage at the output. Assume $|V_{BE}| = 0.7$ V, $\beta = 100$, and neglect the Early effect.
(b) Find the input resistance and the output resistance.
(c) Use the current-gain method to evaluate the voltage gain v_o/v_i.

9.125 For the current mirror in Fig. P9.125, replace the transistors with their hybrid-π models and show that:

$$R_i = \frac{1}{g_{m1}} \parallel r_{o1}$$

$$A_{is} \simeq A_{is}\Big|_{\text{ideal}}\left(1 - \frac{1}{g_{m1} r_{o1}}\right)$$

$$A_{is}\Big|_{\text{ideal}} = g_{m2}/g_{m1}$$

$$R_o = r_{o2}$$

where A_{is} denotes the short-circuit current gain.

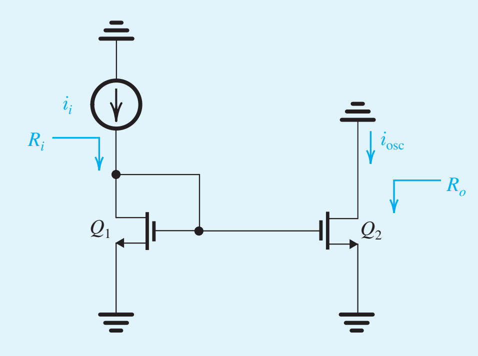


Figure P9.125

****9.126** The MOS differential amplifier shown in Fig. P9.126 utilizes three current mirrors for signal transmission: Q_4–Q_6 has a transmission factor of 2 [i.e., $(W/L)_6/(W/L)_4 = 2$], Q_3–Q_5 has a transmission factor of 1, and Q_7–Q_8 has a transmission factor of 2. All transistors are sized to operate at the same overdrive voltage, $|V_{OV}|$. All transistors have the same Early voltage $|V_A|$.

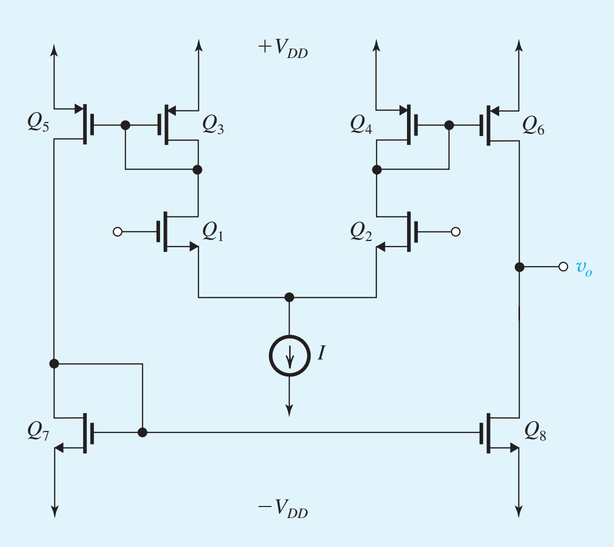


Figure P9.126

(a) Provide in tabular form the values of I_D, g_m, and r_o of each of the eight transistors in terms of I, V_{OV}, and V_A.

(b) Show that the differential voltage gain A_d is given by

$$A_d = 2g_{m1}(r_{o6} \| r_{o8}) = V_A/V_{OV}$$

(c) Show that the CM gain is given by

$$|A_{cm}| \simeq \frac{r_{o6} \| r_{o8}}{R_{SS}} \frac{1}{g_{m7}r_{o7}}$$

where R_{SS} is the output resistance of the bias current source I. [*Hint*: Replace each of Q_1 and Q_2 together with their source resistance $2R_{SS}$ with a controlled current-source $v_{icm}/2R_{SS}$ and an output resistance. For each current mirror, the current transfer ratio is given by

$$A_i \simeq A_i \text{ (ideal)} \left(1 - \frac{1}{g_m r_o}\right)$$

where g_m and r_o are the parameters of the input transistor of the mirror. (see Problem 9.125 above.)]

(d) If the current source I is implemented using a simple mirror and the MOS transistor is operated at the same V_{OV}, show that the CMRR is given by

$$\text{CMRR} = 4(V_A/V_{OV})^2$$

(e) Find the input CM range and the output linear range in terms of V_{DD}, $|V_t|$, and $|V_{OV}|$.

D *****9.127 For the circuit shown in Fig. P9.127, which uses a folded cascode involving transistor Q_3, all transistors have $|V_{BE}| = 0.7$ V for the currents involved, $V_A = 200$ V, and $\beta = 100$. The circuit is relatively conventional except for Q_5, which operates in a Class B mode (we will study this in Chapter 12) to provide an increased negative output swing for low-resistance loads.

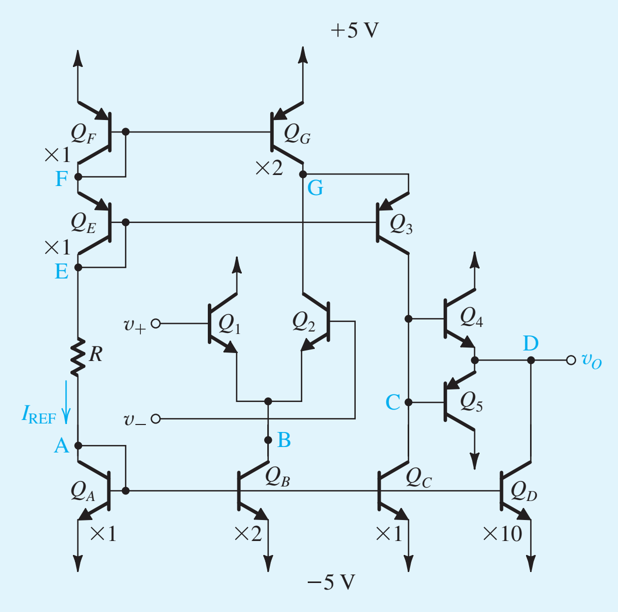


Figure P9.127

(a) Perform a bias calculation assuming $|V_{BE}| = 0.7$ V, high β, $V_A = \infty$, $v_+ = v_- = 0$ V, and v_O is stabilized by feedback to about 0 V. Find R so that the reference current I_{REF} is 100 μA. What are the voltages at all the labeled nodes?

(b) Provide in tabular form the bias currents in all transistors together with g_m and r_o for the signal transistors (Q_1, Q_2, Q_3, Q_4, and Q_5) and r_o for Q_C, Q_D, and Q_G.

(c) Now, using $\beta = 100$, find the voltage gain $v_o/(v_+ - v_-)$, and in the process, verify the polarity of the input terminals.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

(d) Find the input and output resistances.
(e) Find the input common-mode range for linear operation.
(f) For no load, what is the range of available output voltages, assuming $|V_{CEsat}| = 0.3$ V?
(g) Now consider the situation with a load resistance connected from the output to ground. At the positive and negative limits of the output signal swing, find the smallest load resistance that can be driven if one or the other of Q_1 or Q_2 is allowed to cut off.
(g) For what load resistance connected to ground is the output negative voltage limited to −1 V before Q_7 begins to conduct?
(h) For a load resistance one-tenth of that found in (g), what is the output signal swing?

D ***9.128 In the CMOS op amp shown in Fig. P9.128, all MOS devices have $|V_t| = 1$V, $\mu_n C_{ox} = 2\ \mu_p C_{ox} = 40\ \mu$A/V^2, $|V_A| = 50$ V, and $L = 5\ \mu$m. Device widths are indicated on the diagram as multiples of W, where $W = 5\ \mu$m.

(a) Design R to provide a 10-μA reference current.
(b) Assuming $v_O = 0$ V, as established by external feedback, perform a bias analysis, finding all the labeled node voltages, and V_{GS} and I_D for all transistors.
(c) Provide in table form I_D, V_{GS}, g_m, and r_o for all devices.
(d) Calculate the voltage gain $v_o/(v_+ - v_-)$, the input resistance, and the output resistance.
(e) What is the input common-mode range?
(f) What is the output signal range for no load?

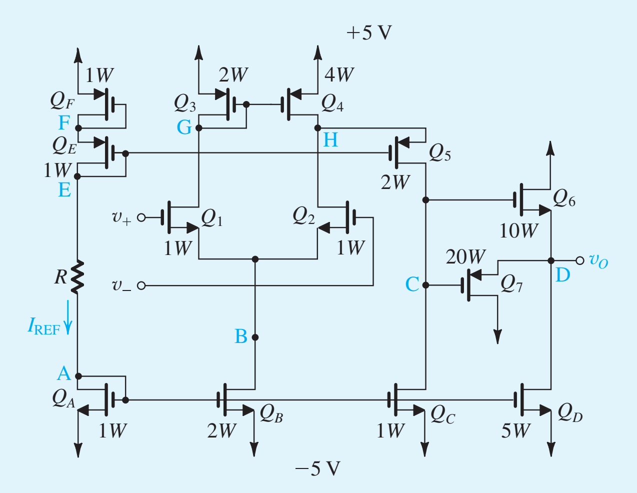


Figure P9.128

CHAPTER 10

Frequency Response

IN THIS CHAPTER YOU WILL LEARN

1. How coupling and bypass capacitors cause the gain of discrete-circuit amplifiers to fall off at low frequencies, and how to obtain an estimate of the frequency f_L at which the gain decreases by 3 dB below its value at midband.

2. The internal capacitive effects present in the MOSFET and the BJT and how to model these effects by adding capacitances to the hybrid-π or T model of each of the two transistor types.

3. The high-frequency limitation on the gain of the CS and CE amplifiers, and how the gain falloff and the upper 3-dB frequency f_H are mostly determined by the small capacitance between the drain and gate (collector and base).

4. Powerful methods for the analysis of the high-frequency response of amplifier circuits of varying complexity.

5. How the cascode amplifier studied in Chapter 8 can be designed to obtain wider bandwidth than is possible with the CS and CE amplifiers.

6. The high-frequency performance of the source and emitter followers.

7. The high-frequency performance of differential amplifiers.

8. Circuit configurations for obtaining wideband amplification.

Introduction

Our study of transistor amplifiers in Chapters 5 through 9 has assumed that their gain is constant independent of the frequency of the input signal. This would imply that their bandwidth is infinite, which of course is not true! To illustrate, we show in Fig. 10.1 a sketch of the magnitude of the gain versus the frequency of the input signal of a discrete-circuit BJT or MOS amplifier. Observe that there is indeed a wide frequency range over which the gain remains almost constant. This is the useful frequency range of operation for the particular amplifier. Thus far, we have been assuming that our amplifiers are operating in this band, called the middle-frequency band or **midband**. The amplifier is designed so that its midband coincides with the frequency spectrum of the signals it is required to amplify. If this were not the case, the amplifier would distort the frequency spectrum of the input signal, with different components of the input signal being amplified by different amounts.

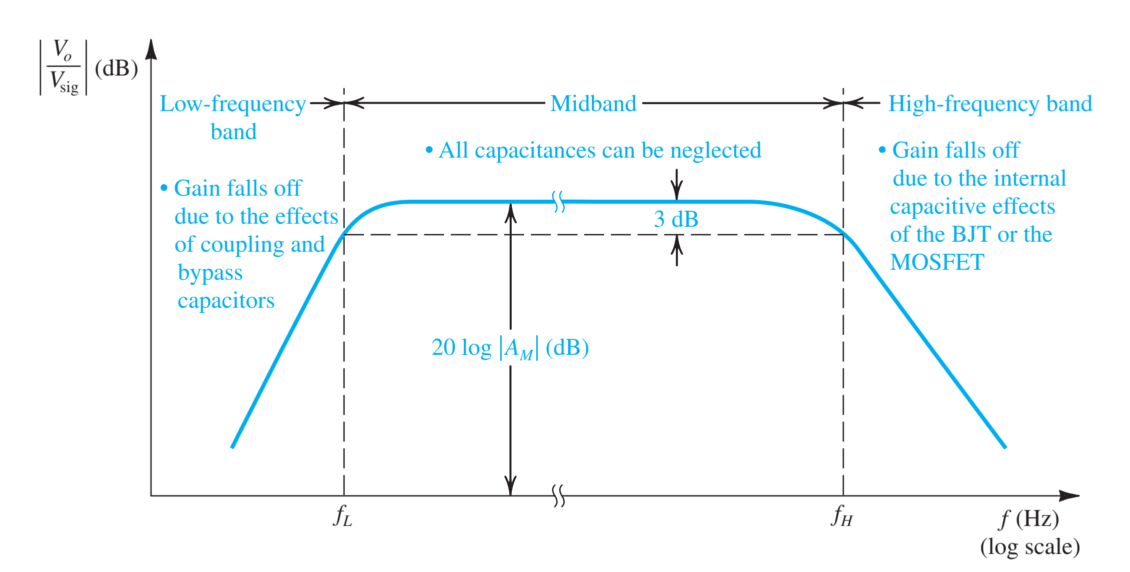


Figure 10.1 Sketch of the magnitude of the gain of a discrete-circuit BJT or MOS amplifier versus frequency. The graph delineates the three frequency bands relevant to frequency-response determination.

Figure 10.1 indicates that at lower frequencies, the magnitude of the amplifier gain falls off. This occurs because the coupling and bypass capacitors no longer have low impedances. Recall that we assumed that their impedances were small enough to act as short circuits. Although this can be true at midband frequencies, as the frequency of the input signal is lowered, the reactance $1/j\omega C$ of each of these capacitors becomes significant and, as will be shown in Section 10.1, this results in a decrease in the overall voltage gain of the amplifier. In the analysis of the low-frequency response of discrete-circuit amplifiers in Section 10.1 we will be particularly interested in the determination of the frequency f_L, which defines the lower end of the midband. It is usually defined as the frequency at which the gain drops by 3 dB below its value in midband. Integrated-circuit amplifiers do not utilize coupling and bypass capacitors, and thus their midband extends down to zero frequency (dc), as shown in Fig. 10.2.

Figures 10.1 and 10.2 indicate that the gain of the amplifier falls off at the high-frequency end. This is due to internal capacitive effects in the BJT and in the MOSFET. We shall study these effects in Section 10.2 and model them with capacitances that we will add to the hybrid-π or T model of the BJT and the MOSFET. The resulting high-frequency device models will be utilized in Section 10.3 in the analysis of the high-frequency response of the CS and CE amplifiers, both discrete and integrated. We will be specifically interested in the determination of the frequency f_H, which defines the upper end of the midband. It is defined as the frequency at which the gain drops by 3 dB below its midband value. Thus, the amplifier bandwidth is defined by f_L and f_H (0 and f_H for IC amplifiers):

$$\text{BW} = f_H - f_L \text{ (discrete-circuit amplifiers)}$$
$$\text{BW} = f_H \text{ (integrated-circuit amplifiers)}$$

A figure of merit for the amplifier is its **gain–bandwidth product**, defined as

$$GB = |A_M|\text{BW}$$

It will be seen that in amplifier design, it is usually possible to trade off gain for bandwidth.

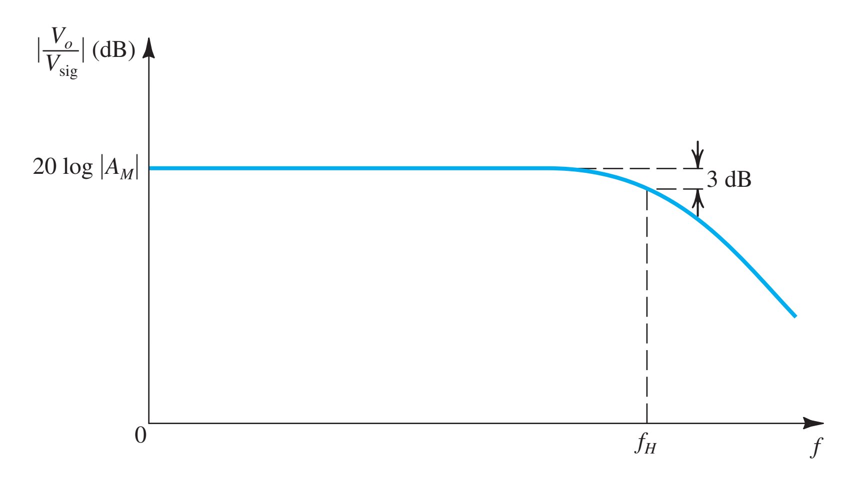


Figure 10.2 Frequency response of a direct-coupled (dc) amplifier. Observe that the gain does *not* fall off at low frequencies, and the midband gain A_M extends down to zero frequency.

The remainder of this chapter will be concerned with the frequency-response analysis of a variety of amplifier configurations of varying degrees of complexity. Of particular interest to us are ways to extend the amplifier bandwidth (i.e., increase f_H) either by adding specific circuit components, such as source and emitter-degeneration resistances, or by changing the circuit configuration altogether.

Before embarking on the study of this chapter, the reader is urged to review Section 1.6, which introduces the subject of amplifier frequency response and the extremely important topic of single-time-constant (STC) circuits. More details on STC circuits can be found in Appendix E. As well, Appendix F provides a review of important tools from circuit and system theory: poles, zeros, and Bode plots.

Finally, a note on notation: Since we will be dealing with quantities that are functions of frequency, or, equivalently, the Laplace variable s, we will be using capital letters with lowercase subscripts for our symbols. This practice conforms with the symbol notation introduced in Chapter 1.

10.1 Low-Frequency Response of Discrete-Circuit Common-Source and Common-Emitter Amplifiers

In this section, we consider the effect of the coupling and bypass capacitors on the gain of discrete-circuit common-source (CS) and common-emitter (CE) amplifiers. As mentioned earlier, their effect manifests itself only at low frequencies (i.e., below the midband). We consider first the CS amplifier, since the infinite input resistance at the gate of the MOSFET makes the analysis of this circuit simpler than that of its CE counterpart.

10.1.1 The CS Amplifier

Figure 10.3(a) shows a discrete-circuit common-source amplifier utilizing the classical biasing arrangement (Section 7.5.1). Two coupling capacitors, C_{C1} and C_{C2}, and a bypass capacitor C_S are employed. At midband frequencies, these large capacitances have negligibly small

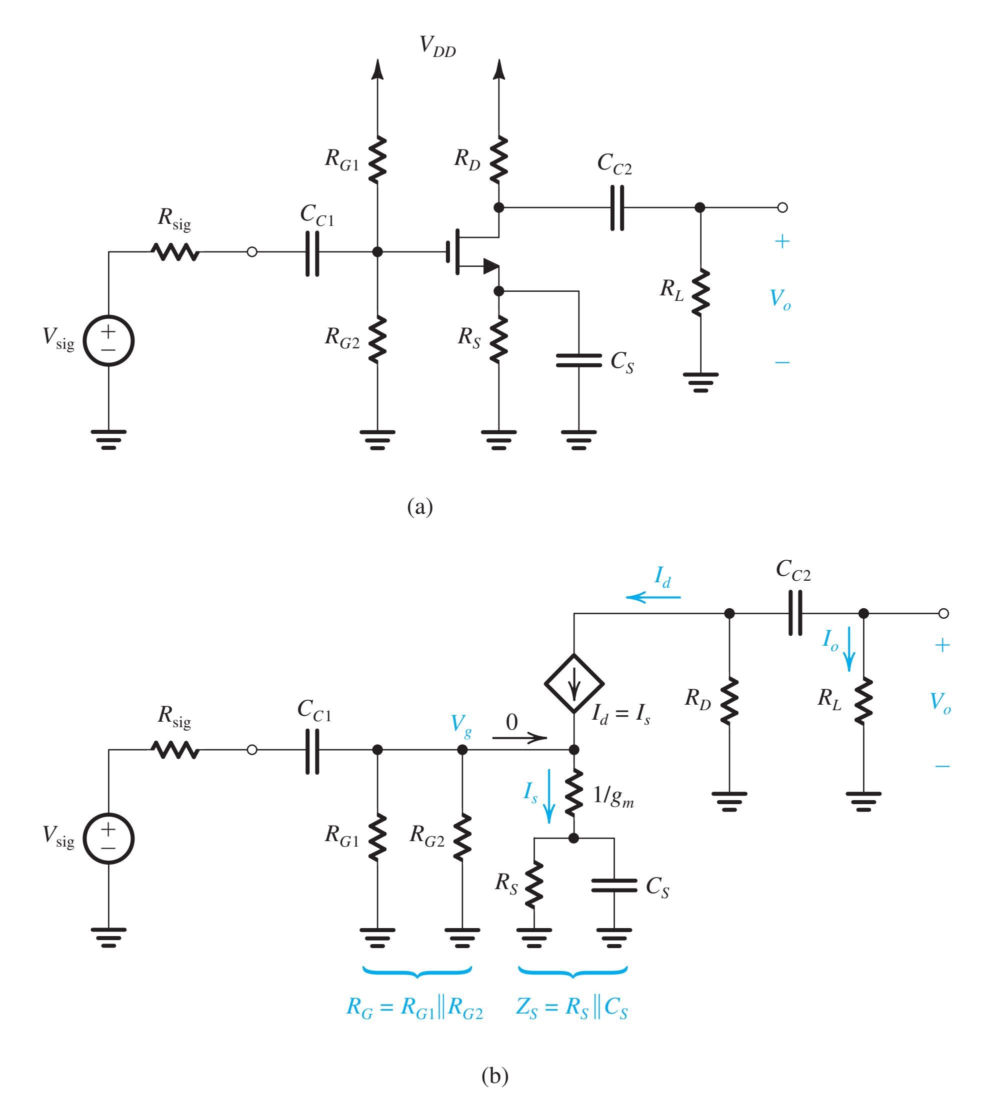


Figure 10.3 (**a**) Capacitatively coupled common-source amplifier. (**b**) The amplifier equivalent circuit at low frequencies. Note that the T model is used for the MOSFET and r_o is neglected.

impedances and can be assumed to be perfect short circuits for the purpose of calculating the midband gain, as was done in Section 7.5.1. However, at low frequencies, the reactance $1/j\omega C$ of each of the three capacitances increases and the amplifier gain decreases, as we shall now show.

To determine the amplifier gain V_o/V_{sig} at low frequencies, we utilize the amplifier equivalent circuit shown in Fig. 10.3(b). This circuit is obtained by short-circuiting V_{DD} and replacing the MOSFET with its T model, which is the most convenient model to use when an impedance, such as Z_S, is present in the source lead. The transistor r_o has been omitted because including it would complicate the analysis considerably and, moreover, as mentioned in Chapter 7, r_o has a negligible effect on the performance of discrete-circuit amplifiers, as can be verified using circuit simulation.

The gain V_o/V_{sig} of the amplifier can be obtained from the equivalent circuit in Fig. 10.3(b) by starting at the source and working our way to the load, as

$$\frac{V_o}{V_{sig}} = \frac{V_g}{V_{sig}} \times \frac{I_d}{V_g} \times \frac{V_o}{I_d}$$

where V_g is the voltage between gate and ground, and I_d is the drain current. To find the fraction of V_{sig} that appears at the transistor gate, V_g, we note that the input resistance at the gate is infinite and thus the amplifier input resistance is $R_G = R_{G1} \| R_{G2}$. Using the voltage-divider rule gives

$$V_g = V_{sig} \frac{R_G}{R_G + \dfrac{1}{sC_{C1}} + R_{sig}}$$

which can be rearranged in the form

$$\frac{V_g}{V_{sig}} = \frac{R_G}{R_G + R_{sig}} \frac{s}{s + \dfrac{1}{C_{C1}(R_G + R_{sig})}} \tag{10.1}$$

Thus, we see that the effect of C_{C1} is to cause the expression for the signal transmission from the signal source to the amplifier input to acquire a frequency-dependent factor. From Section 1.6 we recognize this factor as the transfer function of a single-time-constant circuit of the high-pass type, with a pole frequency ω_{P1},

$$\omega_{P1} = 1/C_{C1}(R_{sig} + R_G) \tag{10.2}$$

In addition to the pole, C_{C1} introduces a zero at $s = 0$ (dc). This is hardly surprising, since C_{C1} is included in the amplifier circuit because it blocks dc. Figure 10.4 shows a sketch of the magnitude of the frequency-dependent factor in the transfer function of Eq. (10.1) versus frequency ω.

Continuing with the analysis, we next determine the drain current I_d, which is equal to the source current I_s. The latter can be found by dividing V_g by the total impedance in the source

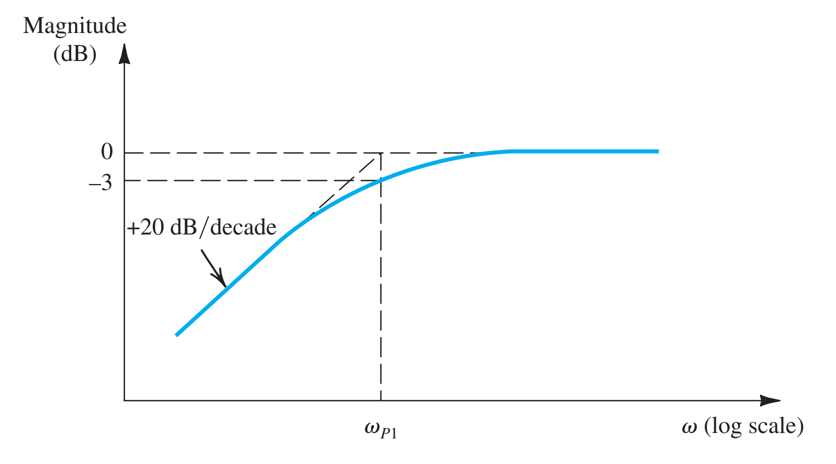


Figure 10.4 Sketch of the magnitude of the high-pass function $\dfrac{s}{s+\omega_{P1}}$, that is, $\dfrac{\omega}{\sqrt{\omega^2 + \omega_{P1}^2}}$ versus frequency ω.

lead, $(1/g_m + Z_S)$,

$$I_d = I_s = \frac{V_g}{\dfrac{1}{g_m} + Z_S} = g_m V_g \frac{Y_S}{g_m + Y_S}$$

where

$$Y_S = \frac{1}{Z_S} = \frac{1}{R_S} + sC_S$$

Thus,

$$\frac{I_d}{V_g} = g_m \frac{s + \dfrac{1}{C_S R_S}}{s + \dfrac{g_m + 1/R_S}{C_S}} \qquad (10.3)$$

Thus, the bypass capacitor introduces a pole with frequency ω_{P2},

$$\omega_{P2} = \frac{g_m + 1/R_S}{C_S} \qquad (10.4)$$

and a transmission zero on the negative real axis of the s plane at

$$s_Z = -\frac{1}{C_S R_S} \qquad (10.5)$$

and thus has a frequency

$$\omega_Z = \frac{1}{C_S R_S} \qquad (10.6)$$

Observe that since g_m is usually large, $\omega_{P2} \gg \omega_Z$. That is, ω_{P2} will be closer to the midband, and thus it plays a more significant role in determining ω_L than does ω_Z. Figure 10.5 shows a sketch of the magnitude of the frequency-dependent factor of the transfer function in Eq. (10.3).

To complete the analysis, we find V_o by first using the current-divider rule to determine the fraction of I_d that flows through R_L,

$$I_o = -I_d \frac{R_D}{R_D + \dfrac{1}{sC_{C2}} + R_L}$$

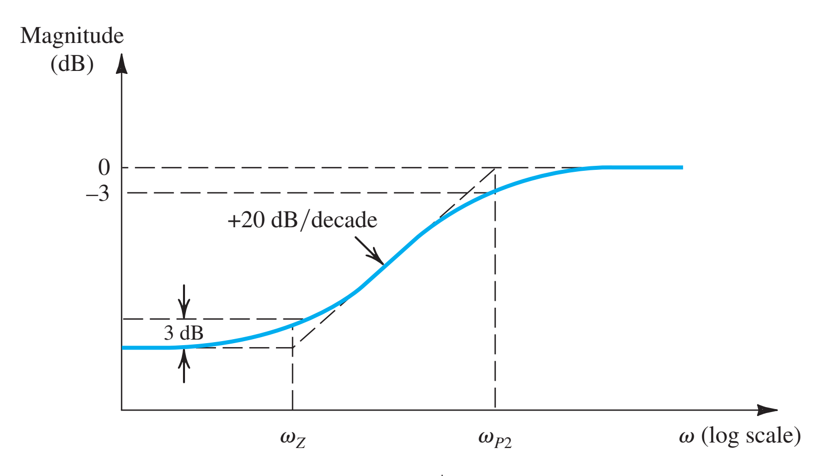


Figure 10.5 Sketch of the magnitude of the function $\dfrac{s + \omega_Z}{s + \omega_{P2}}$ versus frequency ω.

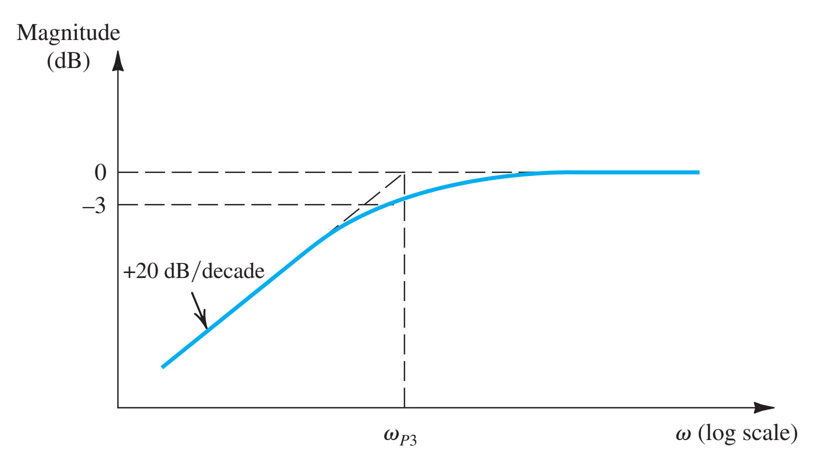


Figure 10.6 Sketch of the magnitude of the high-pass function $\dfrac{s}{s+\omega_{P3}}$ versus frequency ω.

and then multiplying I_o by R_L. The result is

$$\frac{V_o}{I_d} = -\frac{R_D R_L}{R_D + R_L}\,\frac{s}{s + \dfrac{1}{C_{C2}(R_D + R_L)}} \tag{10.7}$$

from which we see that C_{C2} introduces a pole with frequency,

$$\omega_{P3} = \frac{1}{C_{C2}(R_D + R_L)} \tag{10.8}$$

and a zero at $s = 0$ (dc). A sketch of the magnitude of the frequency-dependent factor of the transfer function in Eq. (10.7) is shown in Fig. 10.6.

The overall low-frequency gain function of the amplifier can be found by combining Eqs. (10.1), (10.3), and (10.7),

$$\frac{V_o}{V_{sig}} = -\frac{R_G}{R_G + R_{sig}}\, g_m(R_D \parallel R_L)\left(\frac{s}{s+\omega_{P1}}\right)\left(\frac{s+\omega_Z}{s+\omega_{P2}}\right)\left(\frac{s}{s+\omega_{P3}}\right)$$

$$\frac{V_o}{V_{sig}} = A_M\left(\frac{s}{s+\omega_{P1}}\right)\left(\frac{s+\omega_Z}{s+\omega_{P2}}\right)\left(\frac{s}{s+\omega_{P3}}\right) \tag{10.9}$$

where A_M, the midband gain, is given by

$$A_M = -\frac{R_G}{R_G + R_{sig}}\, g_m(R_D \parallel R_L) \tag{10.10}$$

which is the value we would have obtained, had we assumed that C_{C1}, C_S, and C_{C2} were acting as perfect short circuits. In this regard, note from Eq. (10.9) that at midband frequencies—that is, at frequencies $s = j\omega$ with ω much higher than ω_{P1}, ω_{P2}, ω_Z, and ω_{P3}—V_o/V_{sig} approaches A_M, as should be expected.

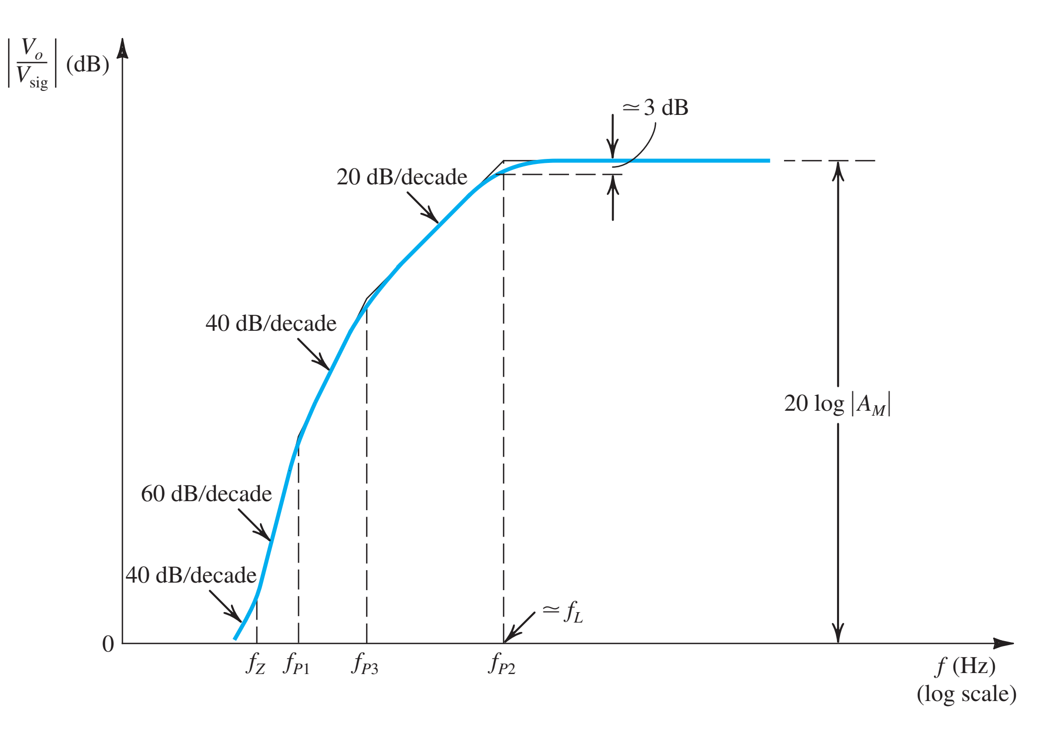


Figure 10.7 Sketch of the low-frequency magnitude response of a CS amplifier for which the three pole frequencies are sufficiently separated for their effects to appear distinct.

Determining the 3-dB Frequency f_L The magnitude of the amplifier gain at a frequency ω can be obtained by substituting $s = j\omega$ in Eq. (10.9) and evaluating the magnitude of the resulting complex function. In this way, the low-frequency response of the amplifier can be plotted versus frequency, and the lower 3-dB frequency f_L can be determined as the frequency at which $|V_o/V_{sig}|$ drops to $|A_M|/\sqrt{2}$.

A simpler approach for determining f_L is possible if the poles and zeros are sufficiently separated from one another. In this case, we can employ the Bode plot rules (see Appendix F) to obtain a Bode plot for the gain magnitude. Such a plot is shown in Fig. 10.7. This graph is simply a combination of the graphs in Figs. 10.4, 10.5, and 10.6. Observe that since the poles and zeros are sufficiently separated, their effects appear distinct. As we move downward in frequency from the midband, we find that at each pole frequency, the slope of the asymptote to the gain function increases by 20 dB/decade, and at the zero frequency (f_Z) it decreases by 20 dB/decade. Note that for the purpose of this sketch, we assumed f_{P2} to be the highest of the three poles and zero frequencies, and that the zero has the lowest frequency.

A quick way for estimating the 3-dB frequency f_L is possible if the highest-frequency pole (here, assumed to be f_{P2}) is separated from the nearest pole or zero (here, f_{P1}) by at least a factor of 4 (two octaves). In such a case, f_L is approximately equal to the highest of the pole frequencies,

$$f_L \simeq f_{P2} \tag{10.11}$$

We refer to this situation as one in which a **dominant pole** exists, with the frequency of the dominant pole being f_{P2}. Of course, if a dominant pole exists, f_L can be estimated without the need for the Bode plot.

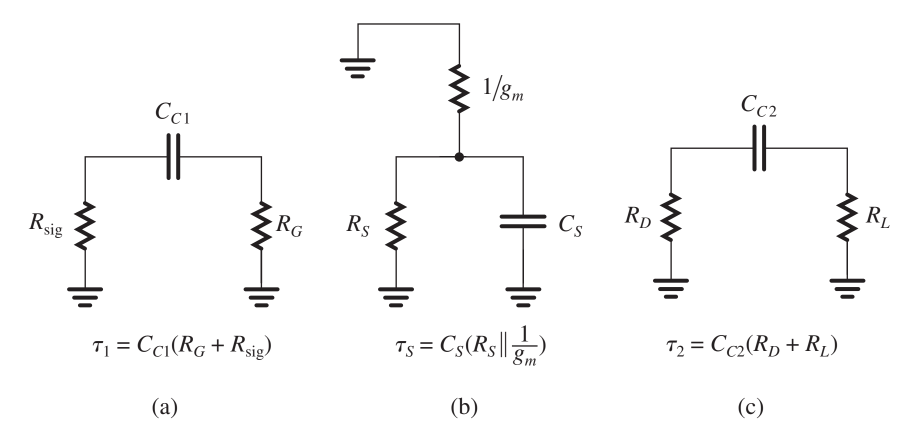


Figure 10.8 Circuits for determining the time constant of each of the three capacitors, and hence the pole associated with each one. Note that this determination is possible because in the circuit of Fig. 10.3, the capacitors do not interact.

If a dominant pole does not exist, the following approximate expression for f_L can be used[1]

$$f_L \simeq \sqrt{f_{P1}^2 + f_{P2}^2 + f_{P3}^2 - 2f_Z^2} \tag{10.12}$$

Determining the Pole and Zero Frequencies by Inspection Since the capacitors in the CS amplifier circuit do not interact, a simple procedure exists for determining the frequencies of the pole and zero introduced by each capacitor.

Consider first the zeros. By its definition, a transmission zero is the value of s at which the input does not reach the output, resulting in $V_o = 0$. Examination of the circuit in Fig. 10.3(b) indicates that C_{C1} becomes an infinite impedance at $s = 0$ and thus introduces a transmission zero at $s = 0$ (i.e., blocks dc).

An identical statement applies to C_{C2}. However, the bypass capacitor C_S has a different effect: Its transmission zero is at the value of s that causes Z_S to become infinite, and hence I_s, I_d, and V_o become zero, which is s_Z given by Eq. (10.5).

To determine the poles, we set $V_{sig} = 0$.[2] This results in the three separate circuits shown in Fig. 10.8. Each of the three circuits can be used to determine the resistance "seen" by the particular capacitor, and hence the time constant associated with this capacitor. The corresponding pole frequency ω_P is the inverse of the time constant.

Selecting Values for the Coupling and Bypass Capacitors We now address the design issue of selecting appropriate values for C_{C1}, C_S, and C_{C2}. The design objective is to place the lower 3-dB frequency f_L at a specified value while minimizing the capacitor values. Since the resistance seen by C_S, $\left(\frac{1}{g_m} \| R_S\right)$, is usually the smallest of the three resistances, the total capacitance is minimized by selecting C_S to provide the highest frequency pole; that is, making its pole frequency $f_{P2} = f_L$. We then decide on the location of the other two pole frequencies, say, 5 to 10 times lower than the frequency of the dominant pole, f_{P2}. However,

---

[1]The derivation of this expression is simple and is given in Chapter 9 of the fourth edition of this book.

[2]The poles of a circuit are its natural modes, and thus are independent of the value of the input signal.

the values selected for f_{P1} and f_{P3} should not be too low, for that would require larger values for C_{C1} and C_{C2} than may be necessary. The design procedure will be illustrated by an example.

Example 10.1

We wish to select appropriate values for the coupling capacitors C_{C1} and C_{C2} and the bypass capacitor C_S for a CS amplifier for which $R_G = 4.7\ \text{M}\Omega$, $R_D = R_L = 15\ \text{k}\Omega$, $R_{sig} = 100\ \text{k}\Omega$, $R_S = 10\ \text{k}\Omega$, and $g_m = 1\ \text{mA/V}$. It is required to have f_L at 100 Hz and that the nearest break frequency be at least a decade lower.

Solution

We select C_S so that

$$f_{P2} = \frac{g_m + \dfrac{1}{R_S}}{2\pi C_S} = f_L$$

Thus,

$$C_S = \frac{1.1 \times 10^{-3}}{2\pi \times 100} = 1.75\ \mu\text{F}$$

For $f_{P1} = f_{P3} = 10$ Hz, we obtain

$$10 = \frac{1}{2\pi C_{C1}(0.1 + 4.7) \times 10^6}$$

which yields

$$C_{C1} = 3.3\ \text{nF}$$

and

$$10 = \frac{1}{2\pi C_{C2}(15 + 15) \times 10^3}$$

which results in

$$C_{C2} = 0.53\ \mu\text{F}$$

Finally, we calculate the frequency of the zero f_Z as

$$f_Z = \frac{1}{2\pi C_S R_S} = \frac{1}{2\pi \times 1.75 \times 10^{-6} \times 10 \times 10^3}$$
$$= 9.1\ \text{Hz}$$

EXERCISE

10.1 A CS amplifier has $C_{C1} = C_S = C_{C2} = 1\ \mu\text{F}$, $R_G = 10\ \text{M}\Omega$, $R_{sig} = 100\,\text{k}\Omega$, $g_m = 2\ \text{mA/V}$, $R_D = R_L = R_S = 10\,\text{k}\Omega$. Find $A_M, f_{P1}, f_{P2}, f_{P3}, f_Z$, and f_L.
Ans. –9.9 V/V; 0.016 Hz; 334.2 Hz; 8 Hz; 15.91 Hz; 334.2 Hz

10.1.2 The Method of Short-Circuit Time Constants

In some circuits, such as that of the common-emitter amplifier discussed shortly, the capacitors interact, making it difficult to determine the pole frequencies. Fortunately, however, there is a simple method for obtaining an estimate for f_L without the need to determine the frequencies of the poles. Although the method is predicated on the assumption that one of the poles is dominant, the resulting estimate for f_L is usually very good even if this assumption is not strictly valid. The method is as follows[3]:

1. Set the input signal $V_{\text{sig}} = 0$.
2. Consider the capacitors one at a time. That is, while considering capacitor C_i, set all the other capacitors to infinite values (i.e., replace them with short circuits—hence the name of the method).
3. For each capacitor C_i, find the total resistance R_i seen by C_i. This can be determined either by inspection or by replacing C_i with a voltage source V_x and finding the current I_x drawn from V_x; $R_i \equiv V_x/I_x$.
4. Calculate the 3-dB frequency f_L using

$$f_L \simeq \sum_{i=1}^{n} \frac{1}{C_i R_i} \tag{10.13}$$

where n is the total number of capacitors.

Besides its simplicity, this method has a very important side benefit: Equation (10.13) indicates the relative contribution of each capacitor to the value of f_L. Specifically, the lower the value of the time constant associated with a particular capacitor, the greater the contribution of this capacitor to f_L. As will be seen shortly, this observation has important design implications. Application of the method of short-circuit time constants will be illustrated in the next section, where it is utilized to determine f_L of the CE amplifier.

10.1.3 The CE Amplifier

Figure 10.9(a) shows a discrete-circuit common-emitter amplifier utilizing the classical biasing arrangement (Section 7.5.2), together with coupling capacitors C_{C1} and C_{C2}, and bypass capacitor C_E. We wish to obtain an estimate of the frequency f_L at which the gain of this amplifier drops by 3 dB below its value at midband. As well, we need to determine how to select appropriate values for C_{C1}, C_E, and C_{C2} to ensure that f_L is placed at a desired location while minimizing the total capacitance value required.

To analyze the low-frequency gain of the CE amplifier, we utilize the equivalent circuit shown in Fig. 10.9(b). This equivalent circuit is obtained by short-circuiting V_{CC} and replacing the BJT with its T model, while neglecting r_o. The decision to neglect r_o is based on the insignificant effect of the transistor's output resistance on the gain of discrete-circuit amplifiers, and the considerable complication its inclusion causes to the analysis. From the circuit in Fig. 10.9(b), we observe that the finite input current in the base of the BJT causes C_{C1} and C_{C2} to interact. That is, unlike the case of the CS amplifier, here each of the two poles caused by C_{C1} and C_{C2} will depend on both capacitor values in a complicated fashion that hinders design insight. Therefore, we shall not attempt to determine the pole frequencies

[3]A proof can be found in Gray and Searle, 1969. (See bibliography in Appendix I.)

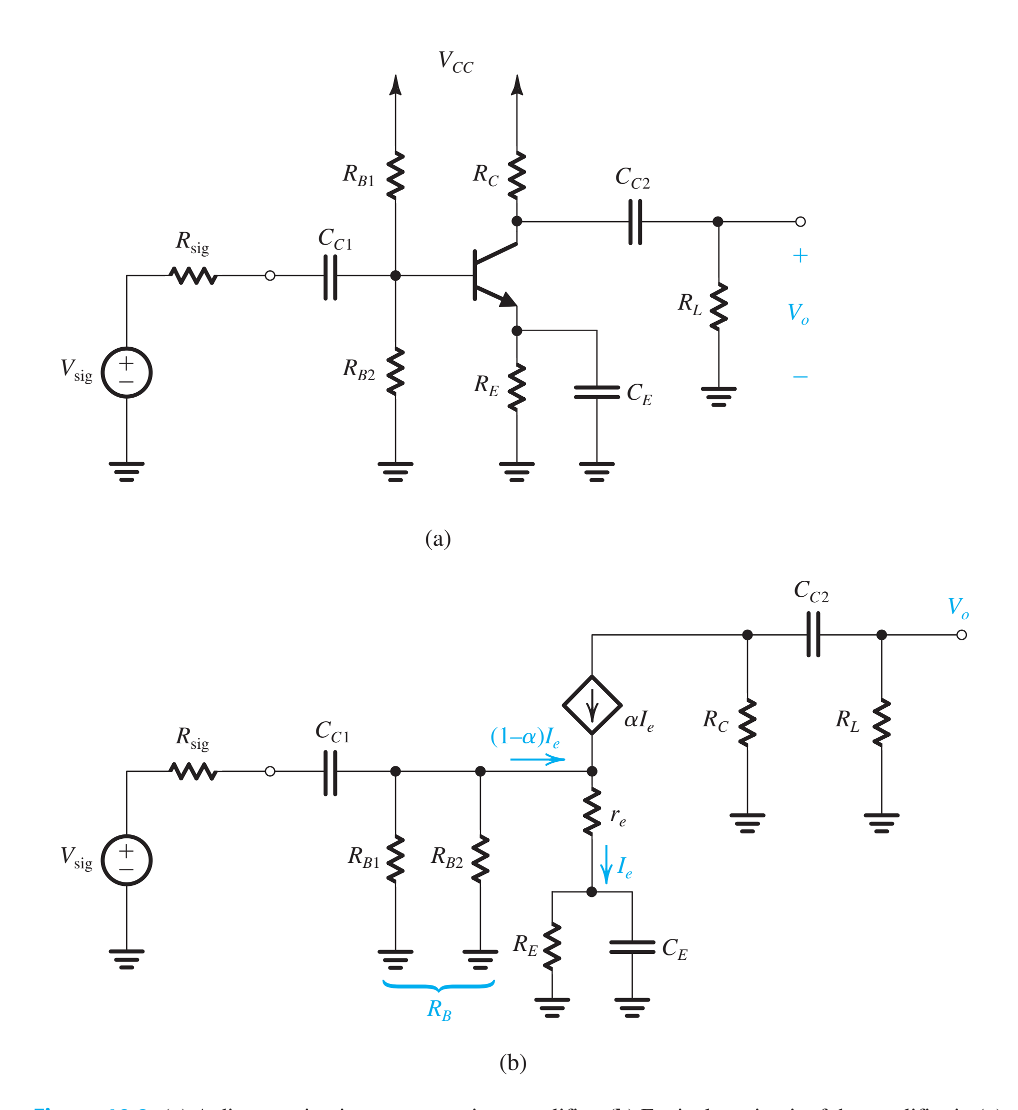


Figure 10.9 (**a**) A discrete-circuit common-emitter amplifier. (**b**) Equivalent circuit of the amplifier in (**a**).

and, instead, we will use the method of short-circuit time constants to obtain an estimate of f_L directly.

Applying the Method of Short-Circuit Time Constants Setting $V_{sig} = 0$ in the circuit of Fig. 10.9(b) and considering each capacitor, one at a time, while short-circuiting the other two results in the three circuits shown in Fig. 10.10. These circuits can be used to determine the resistance seen by each capacitor and hence its effective time constant. For C_{C1} we use the circuit in Fig. 10.10(a) and note that r_π is the input resistance at the base when C_E is short-circuited. Capacitor C_{C1} sees a resistance R_{C1}, which can be found by inspection as

$$R_{C1} = (R_B \| r_\pi) + R_{sig} \tag{10.14}$$

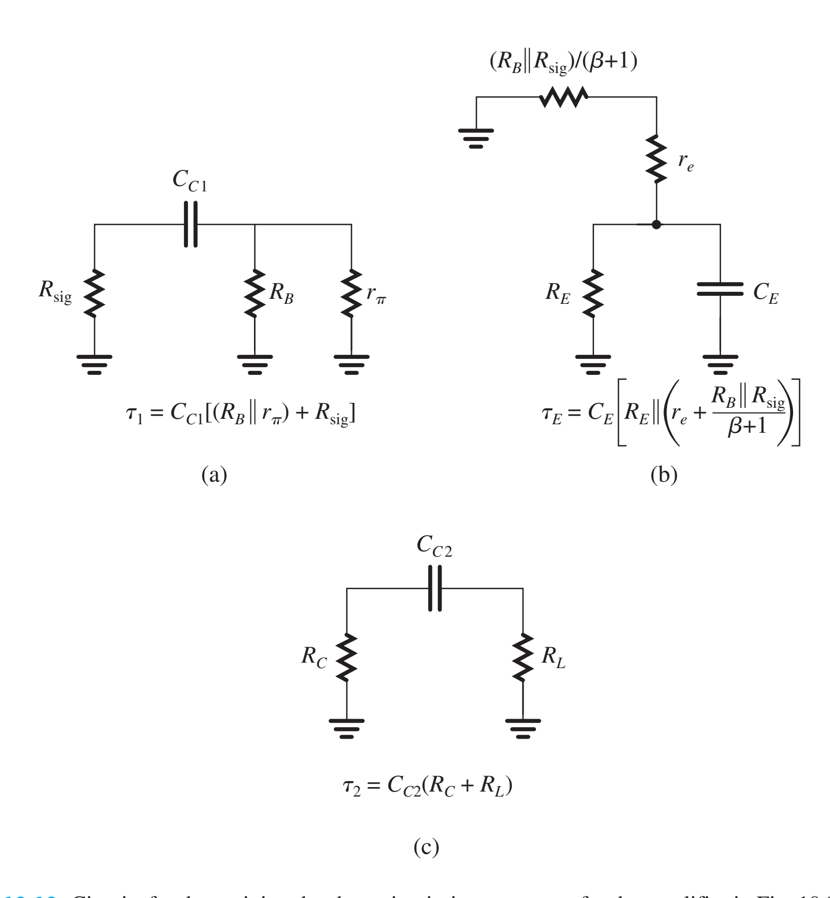


Figure 10.10 Circuits for determining the short-circuit time constants for the amplifier in Fig. 10.9.

and the time constant associated with C_{C1} becomes

$$\tau_{C1} = C_{C1} R_{C1} \tag{10.15}$$

For C_E, we use the circuit in Fig. 10.10(b). Here, we see that with C_{C1} shorted [refer to Fig. 10.9(b)], the resistance in the base becomes $(R_B \| R_{sig})$, which can be reflected to the emitter side as $(R_B \| R_{sig})/(\beta + 1)$. The total resistance R_{CE} seen by C_E can be found by inspection from the circuit in Fig. 10.10(b) as

$$R_{CE} = R_E \| \left[r_e + \frac{R_B \| R_{sig}}{\beta + 1} \right] \tag{10.16}$$

and the time constant becomes

$$\tau_{CE} = C_E R_{CE} \tag{10.17}$$

Finally, the resistance seen by C_{C2} can be determined by inspection of the circuit in Fig. 10.10(c) as

$$R_{C2} = R_C + R_L \tag{10.18}$$

and the corresponding time constant τ_{C2} as

$$\tau_{C2} = C_{C2} R_{C2} \tag{10.19}$$

With the three time constants in hand, the 3-dB frequency f_L can be found from

$$f_L = \frac{\omega_L}{2\pi} = \frac{1}{2\pi}\left[\frac{1}{C_{C1}R_{C1}} + \frac{1}{C_E R_E} + \frac{1}{C_{C2}R_{C2}}\right] \tag{10.20}$$

When numerical values are substituted in this expression, it quickly becomes obvious which of the three capacitors is contributing the most to f_L. Obviously, it is the capacitor that has the smallest time constants. In the CE amplifier, this is usually C_E because the associated resistance R_{CE} is typically small. Knowing which of the capacitors has the potential of dominating the determination of f_L has significant design implications, as shown next.

Selecting Values for C_{C1}, C_E, and C_{C2} We now address the design issue of selecting appropriate values for C_{C1}, C_E, and C_{C2}. The design objective is to place the lower 3-dB frequency f_L at a specified location while minimizing the capacitor values. Since, as mentioned above, C_E usually sees the lowest of the three resistances, the total capacitance is minimized by selecting C_E so that its contribution to f_L is dominant. That is, by reference to Eq. (10.20) we may select C_S so that $1/(C_E R_E)$ is, say, 80% of $\omega_L = 2\pi f_L$, leaving each of the other capacitors to contribute 10% to the value of ω_L. Example 10.2 should help illustrate this process.

Example 10.2

We wish to select appropriate values for C_{C1}, C_{C2}, and C_E for the common-emitter amplifier, which has $R_B = 100$ kΩ, $R_C = 8$ kΩ, $R_L = 5$ kΩ, $R_{sig} = 5$ kΩ, $R_E = 5$ kΩ, $\beta = 100$, $g_m = 40$ mA/V, and $r_\pi = 2.5$ kΩ. It is required to have $f_L = 100$ Hz.

Solution

We first determine the resistances seen by the three capacitors C_{C1}, C_E, and C_{C2} as follows:

$$R_{C1} = (R_B \| r_\pi) + R_{sig}$$
$$= (100 \| 2.5) + 5 = 7.44 \text{ k}\Omega$$
$$R_{CE} = R_E \| \left[r_e + \frac{R_B \| R_{sig}}{\beta + 1}\right]$$
$$= 5 \| \left(0.025 + \frac{100 \| 5}{101}\right) = 0.071 \text{ k}\Omega$$
$$R_{C2} = R_C + R_L = 8 + 5 = 13 \text{ k}\Omega$$

Now, selecting C_E so that it contributes 80% of the value of ω_L gives

$$\frac{1}{C_E \times 71} = 0.8 \times 2\pi \times 100$$
$$C_E = 28 \ \mu\text{F}$$

Next, if C_{C1} is to contribute 10% of f_L,

$$\frac{1}{C_{C1} \times 7.44 \times 10^3} = 0.1 \times 2\pi \times 100$$
$$C_{C1} = 2.1 \ \mu\text{F}$$

Similarly, if C_{C2} is to contribute 10% of f_L, its value should be selected as follows:

$$\frac{1}{C_{C2} \times 13 \times 10^3} = 0.1 \times 2\pi \times 100$$
$$C_{C2} = 1.2\ \mu\text{F}$$

In practice, we would select the nearest standard values for the three capacitors while ensuring that $f_L \leq 100$ Hz. Finally, the frequency of the zero introduced by C_E can be found,

$$f_Z = \frac{1}{2\pi C_E R_E} = \frac{1}{2\pi \times 28 \times 10^{-6} \times 5 \times 10^3} = 1.1\ \text{Hz}$$

which is very far from f_L and thus has an insignificant effect.

EXERCISE

10.2 A common-emitter amplifier has $C_{C1} = C_E = C_{C2} = 1\,\mu\text{F}$, $R_B = 100\,\text{k}\Omega$, $R_{\text{sig}} = 5\,\text{k}\Omega$, $g_m = 40$ mA/V, $r_\pi = 2.5\,\text{k}\Omega$, $R_E = 5\,\text{k}\Omega$, $R_C = 8\,\text{k}\Omega$, and $R_L = 5\,\text{k}\Omega$. Find the value of the time constant associated with each capacitor, and hence estimate the value of f_L. Also compute the frequency of the transmission zero introduced by C_E and comment on its effect on f_L.
Ans. $\tau_{C1} = 7.44$ ms; $\tau_{CE} = 0.071$ ms; $\tau_{C2} = 13$ ms; $f_L = 2.28$ kHz; $f_Z = 31.8$ Hz, which is much smaller than f_L and thus has a negligible effect on f_L

10.2 Internal Capacitive Effects and the High-Frequency Model of the MOSFET and the BJT

While coupling and bypass capacitors cause the gain of transistor amplifiers to fall off at the low-frequency end, the gain falloff at high frequencies is caused by the capacitive effects internal to the transistors. In this section we shall briefly consider these effects and, more importantly, show how the device small-signal model can be augmented to take these effects into account.

10.2.1 The MOSFET

From our study of the physical operation of the MOSFET in Section 5.1, we know that the device has internal capacitances. In fact, we used one of these, the gate-to-channel capacitance, in our derivation of the MOSFET *i*–*v* characteristics. We did, however, implicitly assume that the steady-state charges on these capacitances are acquired instantaneously. In other words,

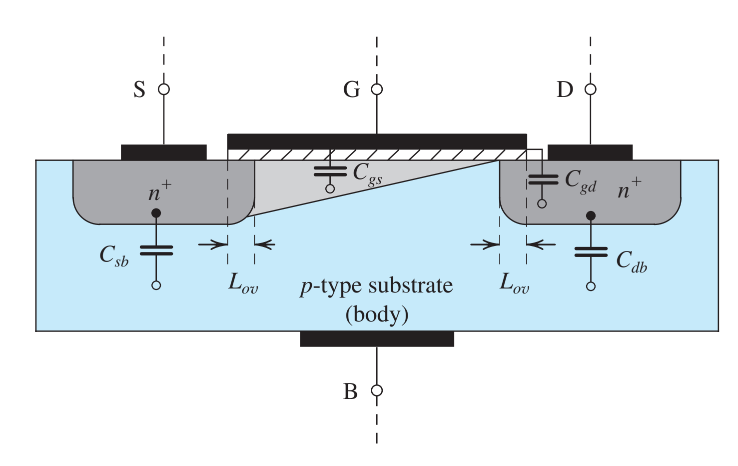


Figure 10.11 A cross section of the n-channel MOSFET operating in the saturation region. The four internal capacitances, C_{gs}, C_{gd}, C_{sb}, and C_{db}, are indicated. Note that the bias voltages are not shown. Also not shown, to keep the diagram simple, is the depletion region.

we did not account for the finite time required to charge and discharge the various internal capacitances. As a result, the device models we derived, such as the small-signal model, do not include any capacitances. The use of these models would predict constant amplifier gains independent of frequency. We know, however, that this (unfortunately) does not happen; in fact, the gain of every MOSFET amplifier falls off at some high frequency. Similarly, the MOSFET digital logic inverter (Chapter 14) exhibits a finite nonzero propagation delay. To be able to predict these results, the MOSFET model must be augmented by including internal capacitances. This is the subject of this section.

To visualize the physical origin of the various internal capacitances, refer to Fig. 10.11, which shows the cross section of an n-channel MOSFET operating in the saturation region, as signified by the tapered n channel that is pinched off at the drain end. As indicated, there are four internal capacitances: Two of these, C_{gs} and C_{gd}, result from the gate-capacitance effect; the other two, C_{sb} and C_{db}, are the depletion capacitances of the pn junctions formed by the source region and the substrate, and the drain region and the substrate, respectively.

The gate-capacitive effect was discussed in Section 5.1. Briefly, the polysilicon gate forms a parallel-plate capacitor with the channel region, with the oxide layer serving as the capacitor dielectric. The gate (or oxide) capacitance per unit gate area is denoted C_{ox}. When the channel is tapered and pinched off, the gate capacitance is given by $\frac{2}{3}WLC_{ox}$. In addition to this capacitance, there are two other small capacitances resulting from the overlap of the gate with the source region (or source diffusion) and the drain region (or drain diffusion). Each of these overlaps has a length L_{ov} and thus the resulting overlap capacitances C_{ov} are given by

$$C_{ov} = WL_{ov}C_{ox} \tag{10.21}$$

Typically, $L_{ov} = 0.05$ to $0.1L$. We can now express the gate-to-source capacitance C_{gs} as

$$C_{gs} = \frac{2}{3}WLC_{ox} + C_{ov} \tag{10.22}$$

For the gate-to-drain capacitance, we note that the channel pinch-off at the drain end causes C_{gd} to consist entirely of the overlap component C_{ov},

$$C_{gd} = C_{ov} \tag{10.23}$$

The depletion-layer capacitances of the two reverse-biased *pn* junctions formed between each of the source and the drain diffusions and the *p*-type substrate (body) can be determined using the formula developed in Section 3.6 (Eq. 3.47). Thus, for the source diffusion, we have the source-body capacitance, C_{sb},

$$C_{sb} = \frac{C_{sb0}}{\sqrt{1 + \frac{V_{SB}}{V_0}}} \tag{10.24}$$

where C_{sb0} is the value of C_{sb} at zero body-source bias, V_{SB} is the magnitude of the reverse-bias voltage, and V_0 is the junction built-in voltage (0.6 V to 0.8 V). Similarly, for the drain diffusion, we have the drain-body capacitance C_{db},

$$C_{db} = \frac{C_{db0}}{\sqrt{1 + \frac{V_{DB}}{V_0}}} \tag{10.25}$$

where C_{db0} is the capacitance value at zero reverse-bias voltage and V_{DB} is the magnitude of this reverse-bias voltage. Note that we have assumed that for both junctions, the grading coefficient $m = \frac{1}{2}$.

It should be noted also that each of these junction capacitances includes a component arising from the bottom side of the diffusion and a component arising from the *sidewalls* of the diffusion. In this regard, observe that each diffusion has three sidewalls that are in contact with the substrate and thus contribute to the junction capacitance (the fourth wall is in contact with the channel).

The formulas for the junction capacitances in Eqs. (10.24) and (10.25) assume small-signal operation. Typical values for the various capacitances exhibited by an *n*-channel MOSFET in a 0.5-μm CMOS process are given in the following exercise.

EXERCISE

10.3 For an *n*-channel MOSFET with $t_{ox} = 10$ nm, $L = 1.0\ \mu$m, $W = 10\ \mu$m, $L_{ov} = 0.05\ \mu$m, $C_{sb0} = C_{db0} = 10$ fF, $V_0 = 0.6$ V, $V_{SB} = 1$ V, and $V_{DS} = 2$ V, calculate the following capacitances when the transistor is operating in saturation: C_{ox}, C_{ov}, C_{gs}, C_{gd}, C_{sb}, and C_{db}.
Ans. 3.45 fF/μm^2; 1.72 fF; 24.7 fF; 1.72 fF; 6.1 fF; 4.1 fF

The High-Frequency MOSFET Model Figure 10.12(a) shows the small-signal model of the MOSFET, including the four capacitances C_{gs}, C_{gd}, C_{sb}, and C_{db}. This model can be used to predict the high-frequency response of MOSFET amplifiers. It is, however, quite complex for manual analysis, and its use is limited to computer simulation using, for example, SPICE. Fortunately, when the source is connected to the body, the model simplifies considerably, as shown in Fig. 10.12(b). In this model, C_{gd}, although small, plays a significant role in determining the high-frequency response of amplifiers and thus must be kept in the

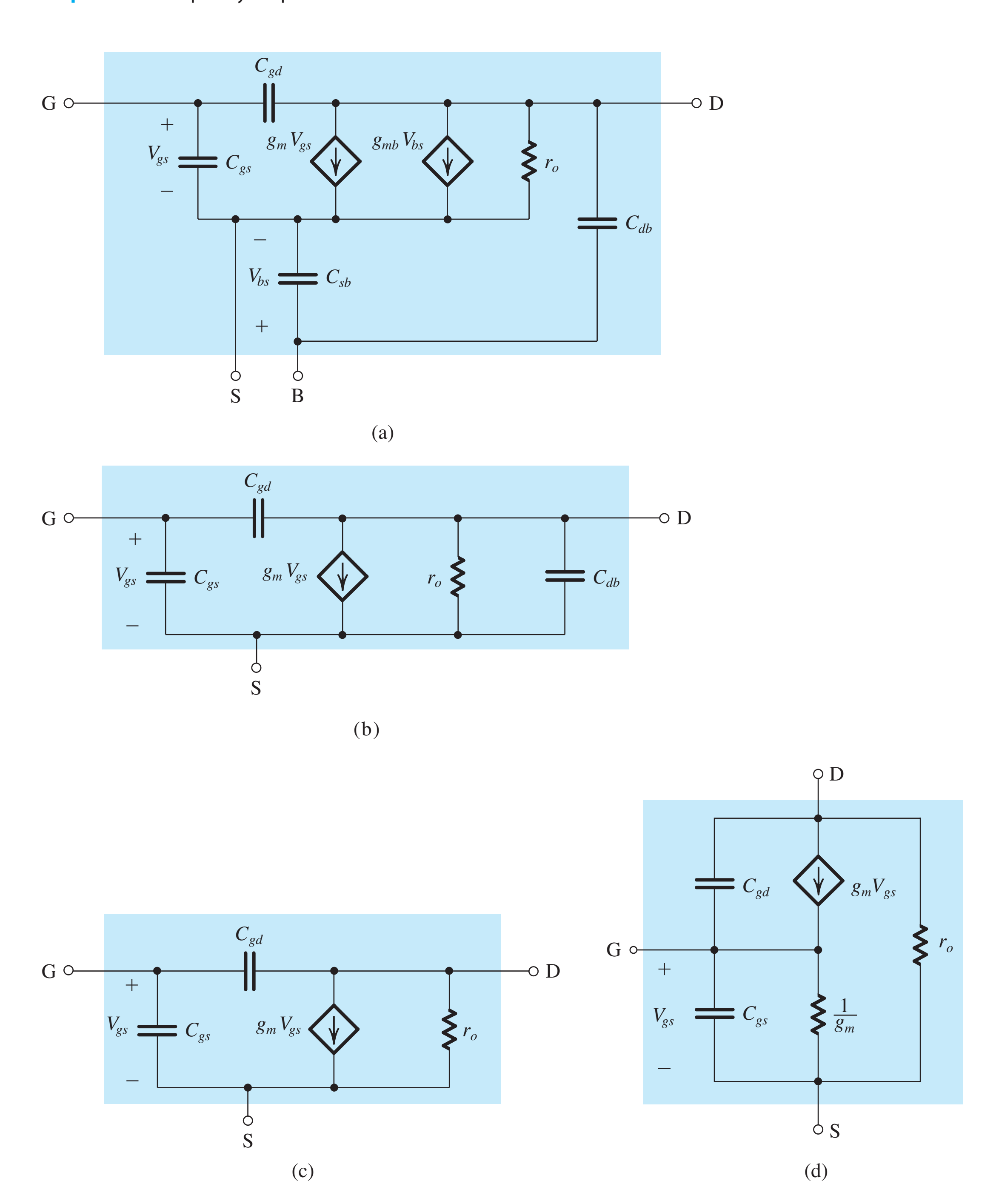


Figure 10.12 (**a**) High-frequency, equivalent-circuit model for the MOSFET. (**b**) The equivalent circuit for the case in which the source is connected to the substrate (body). (**c**) The equivalent-circuit model of (**b**) with C_{db} neglected (to simplify analysis). (**d**) The simplified high-frequency T model.

model. Capacitance C_{db}, on the other hand, can usually be neglected, resulting in significant simplification of manual analysis. The resulting circuit is shown in Fig. 10.12(c). Finally, we show in Fig. 10.12(d) the high-frequency T model in its simplified form.

The MOSFET Unity-Gain Frequency (f_T) A figure of merit for the high-frequency operation of the MOSFET as an amplifier is the unity-gain frequency, f_T, also known as the **transition frequency**, which gives rise to the subscript T. This is defined as *the frequency at which the short-circuit current gain of the common-source configuration becomes unity*. Figure 10.13 shows the MOSFET hybrid-π model with the source as the common terminal between the input and output ports. To determine the short-circuit current gain, the input is fed with a current-source signal I_i and the output terminals are short-circuited. It can be seen that the current in the short circuit is given by

$$I_o = g_m V_{gs} - sC_{gd} V_{gs}$$

Recalling that C_{gd} is small, at the frequencies of interest we can neglect the second term in this equation,

$$I_o \simeq g_m V_{gs} \tag{10.26}$$

From Fig. 10.13, we can express V_{gs} in terms of the input current I_i as

$$V_{gs} = \frac{I_i}{s(C_{gs} + C_{gd})} \tag{10.27}$$

Equations (10.26) and (10.27) can be combined to obtain the short-circuit current gain,

$$\frac{I_o}{I_i} = \frac{g_m}{s(C_{gs} + C_{gd})} \tag{10.28}$$

For physical frequencies $s = j\omega$, it can be seen that the magnitude of the current gain is

$$\left|\frac{I_o}{I_i}\right| = \frac{g_m}{\omega(C_{gs} + C_{gd})}$$

and it becomes unity at the frequency

$$\omega_T = g_m/(C_{gs} + C_{gd})$$

Thus the unity-gain frequency $f_T = \omega_T/2\pi$ is

$$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})} \tag{10.29}$$

Since f_T is proportional to g_m, which determines the midband gain, and inversely proportional to the MOSFET internal capacitances, which limit the amplifier bandwidth, the higher the value of f_T, the more effective the MOSFET becomes as an amplifier. Substituting for g_m using Eq. (7.41), we can express f_T in terms of the bias current I_D (see Problem 10.15). Alternatively, we can substitute for g_m from Eq. (7.40) to express f_T in terms of the overdrive voltage V_{OV} (see Problem 10.16). Both expressions yield additional insight into the high-frequency operation of the MOSFET. The reader is also referred to Appendix G for a further discussion of f_T.

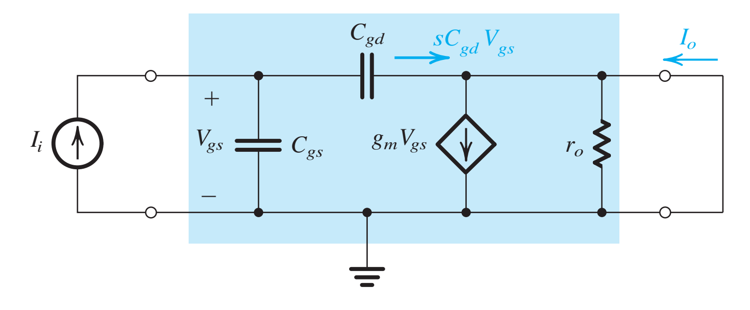


Figure 10.13 Determining the short-circuit current gain I_o/I_i.

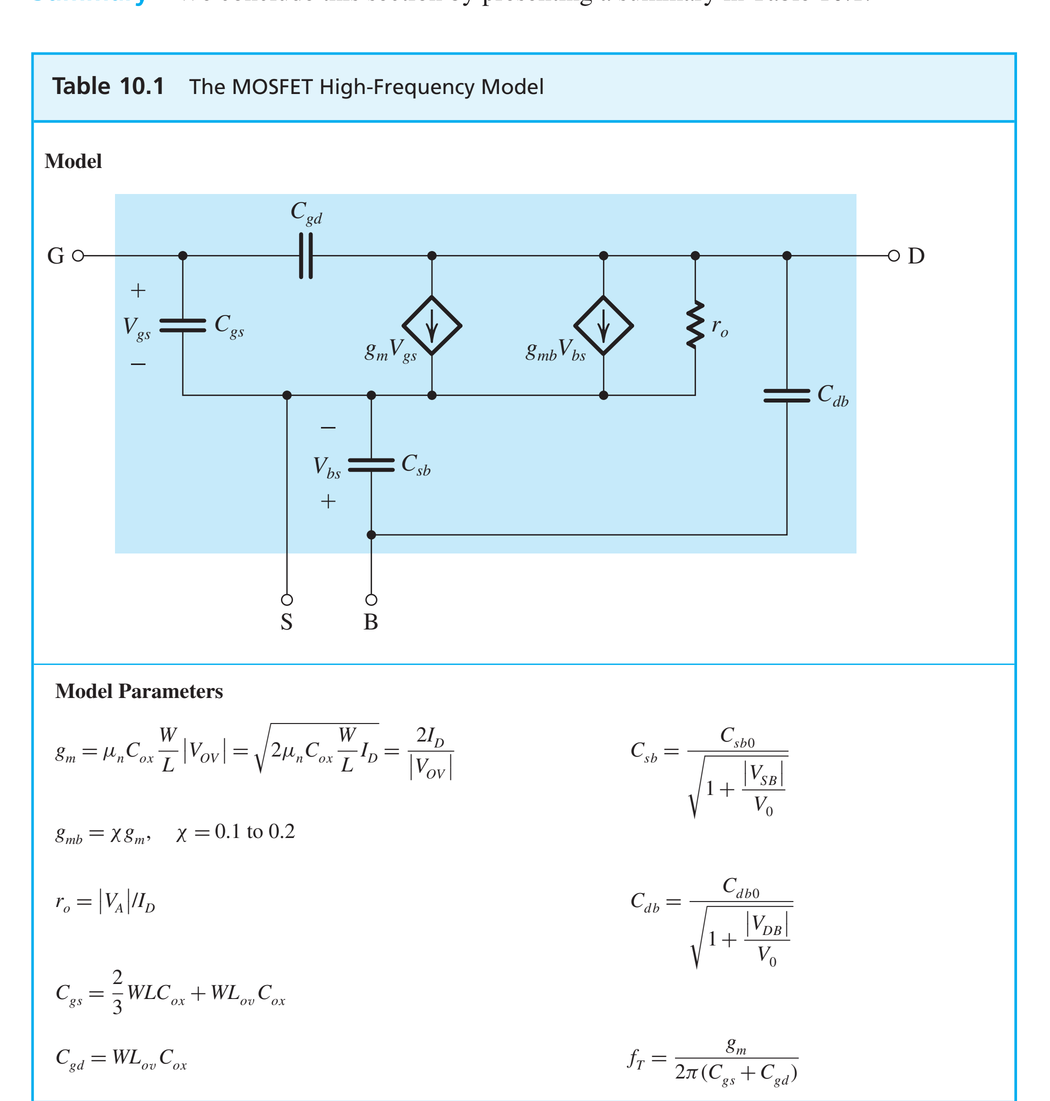
Typically, f_T ranges from about 100 MHz for the older technologies (e.g., a 5-μm CMOS process) to many GHz for newer high-speed technologies (e.g., a 0.13-μm CMOS process).

EXERCISE

10.4 Calculate f_T for the n-channel MOSFET whose capacitances were found in Exercise 10.3. Assume operation at 100 μA and that $k_n' = 160\ \mu\text{A/V}^2$.
Ans. 3.4 GHz

Summary We conclude this section by presenting a summary in Table 10.1.

Table 10.1 The MOSFET High-Frequency Model



Model Parameters

$$g_m = \mu_n C_{ox} \frac{W}{L} |V_{OV}| = \sqrt{2\mu_n C_{ox} \frac{W}{L} I_D} = \frac{2I_D}{|V_{OV}|}$$

$$g_{mb} = \chi g_m, \quad \chi = 0.1 \text{ to } 0.2$$

$$r_o = |V_A|/I_D$$

$$C_{gs} = \frac{2}{3} WLC_{ox} + WL_{ov}C_{ox}$$

$$C_{gd} = WL_{ov}C_{ox}$$

$$C_{sb} = \frac{C_{sb0}}{\sqrt{1 + \frac{|V_{SB}|}{V_0}}}$$

$$C_{db} = \frac{C_{db0}}{\sqrt{1 + \frac{|V_{DB}|}{V_0}}}$$

$$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$$

10.2.2 The BJT

In our study of the physical operation of the BJT in Section 6.1, we assumed transistor action to be instantaneous, and as a result the transistor models we developed do not include any elements (i.e., capacitors or inductors) that would cause time or frequency dependence. Actual transistors, however, exhibit charge-storage phenomena that limit the speed and frequency of their operation. We have already encountered such effects in our study of the *pn* junction in Chapter 3 and have learned that they can be modeled using capacitances. In the following we study the charge-storage effects that take place in the BJT and take them into account by adding capacitances to the hybrid-π and the T models. The resulting augmented BJT model will be able to predict the observed dependence of amplifier gain on frequency, and the time delays that transistor switches and logic gates exhibit.

The Base-Charging or Diffusion Capacitance C_{de} When the transistor is operating in the active mode, minority-carrier charge is stored in the base region. For an *npn* transistor, the stored electron charge in the base, Q_n, can be expressed in terms of the collector current i_C as

$$Q_n = \tau_F i_C \tag{10.30}$$

where τ_F is a device constant with the dimension of time. It is known as the **forward base-transit time** and represents the average time a charge carrier (electron) spends in crossing the base. Typically, τ_F is in the range of 10 ps to 100 ps.

Equation (10.30) applies for large signals and, since i_C is exponentially related to v_{BE}, Q_n will similarly depend on v_{BE}. For small signals, we can define the **small-signal diffusion capacitance C_{de}**,

$$C_{de} \equiv \frac{dQ_n}{dv_{BE}} \tag{10.31}$$

$$= \tau_F \frac{di_C}{dv_{BE}}$$

resulting in

$$C_{de} = \tau_F g_m = \tau_F \frac{I_C}{V_T} \tag{10.32}$$

where I_C is the dc collector bias current at which the transistor is operating. Thus, whenever v_{BE} changes by v_{be}, the collector current changes by $g_m v_{be}$ and the charge stored in the base changes by $C_{de} v_{be} = (\tau_F g_m) v_{be}$. This incremental charge has to be supplied by the base current.

The Base–Emitter Junction Capacitance C_{je} A change in v_{BE} changes not only the charge stored in the base region but also the charge stored in the base–emitter depletion layer. This distinct charge-storage effect is represented by the EBJ depletion-layer capacitance, C_{je}. From the development in Chapter 3, we know that for a forward-biased junction, which the EBJ is, the depletion-layer capacitance is given approximately by

$$C_{je} \simeq 2C_{je0} \tag{10.33}$$

where C_{je0} is the value of C_{je} at zero EBJ voltage.

The Collector–Base Junction Capacitance C_μ In active-mode operation, the CBJ is reverse biased, and its junction or **depletion capacitance**, usually denoted C_μ, can be found from

$$C_\mu = \frac{C_{\mu 0}}{\left(1 + \frac{V_{CB}}{V_{0c}}\right)^m} \tag{10.34}$$

where $C_{\mu 0}$ is the value of C_μ at zero voltage; V_{CB} is the magnitude of the CBJ reverse-bias voltage, V_{0c} is the CBJ built-in voltage (typically, 0.75 V), and m is its grading coefficient (typically, 0.2–0.5).

The High-Frequency Models Figure 10.14 shows the hybrid-π and T models of the BJT, including capacitive effects. Specifically, there are two capacitances: the emitter–base capacitance $C_\pi = C_{de} + C_{je}$ and the collector–base capacitance C_μ. Typically, C_π is in the range of a few picofarads to a few tens of picofarads, and C_μ is in the range of a fraction of a picofarad to a few picofarads.[4] Note that we have also added a resistor r_x to model

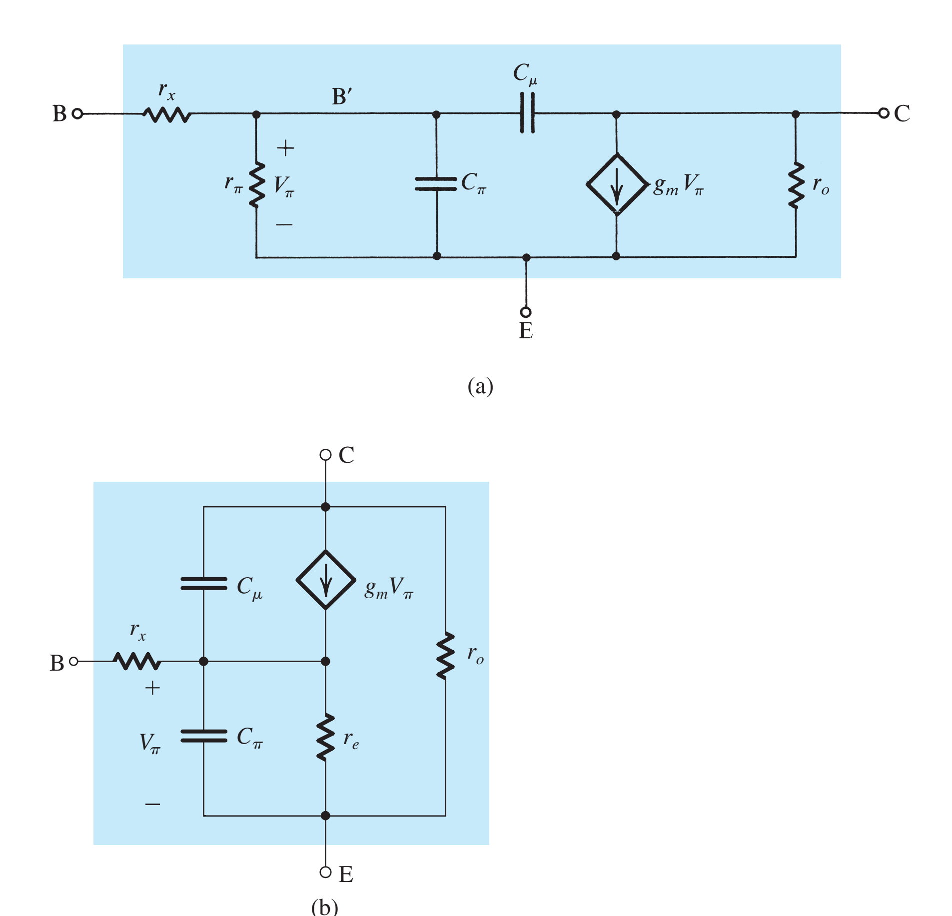


Figure 10.14 The high-frequency models of the BJT: (**a**) hybrid-π model and (**b**) T model.

[4]These values apply for discrete devices and devices fabricated with a relatively old IC process technology (the so-called high-voltage process, see Appendix K). For modern IC fabrication processes, C_π and C_μ are in the range of tens of femtofarads (fF).

the resistance of the silicon material of the base region between the base terminal B and a fictitious internal, or intrinsic, base terminal B′ that is right under the emitter region (refer to Fig. 6.6). Typically, r_x is a few tens of ohms, and its value depends on the current level in a rather complicated manner. Since (usually) $r_x \ll r_\pi$, its effect is negligible at low frequencies. Its presence is felt, however, at high frequencies, as will become apparent later.

The values of the model parameters can be determined at a given bias point using the formulas presented in this section and in Chapter 6. They can also be found from the terminal measurements specified on the BJT data sheets. For computer simulation, SPICE uses the parameters of the given IC technology to evaluate the BJT model parameters (see Appendix B).

The BJT Unity-Gain Frequency The transistor data sheets do not usually specify the value of C_π. Rather, the behavior of β (or h_{fe}) versus frequency is normally given. In order to determine C_π and C_μ, we shall derive an expression for h_{fe}, the CE short-circuit current gain, as a function of frequency in terms of the hybrid-π components. For this purpose consider the circuit shown in Fig. 10.15, in which the collector is shorted to the emitter. A node equation at C provides the short-circuit collector current I_c as

$$I_c = (g_m - sC_\mu)V_\pi \tag{10.35}$$

A relationship between V_π and I_b can be established by multiplying I_b by the impedance seen between B′ and E:

$$V_\pi = I_b(r_\pi \parallel C_\pi \parallel C_\mu) = \frac{I_b}{1/r_\pi + sC_\pi + sC_\mu} \tag{10.36}$$

Thus h_{fe} can be obtained by combining Eqs. (10.35) and (10.36):

$$h_{fe} \equiv \frac{I_c}{I_b} = \frac{g_m - sC_\mu}{1/r_\pi + s(C_\pi + C_\mu)} \tag{10.37}$$

At the frequencies for which this model is valid, $\omega C_\mu \ll g_m$; thus we can neglect the sC_μ term in the numerator and write

$$h_{fe} \simeq \frac{g_m r_\pi}{1 + s(C_\pi + C_\mu)r_\pi}$$

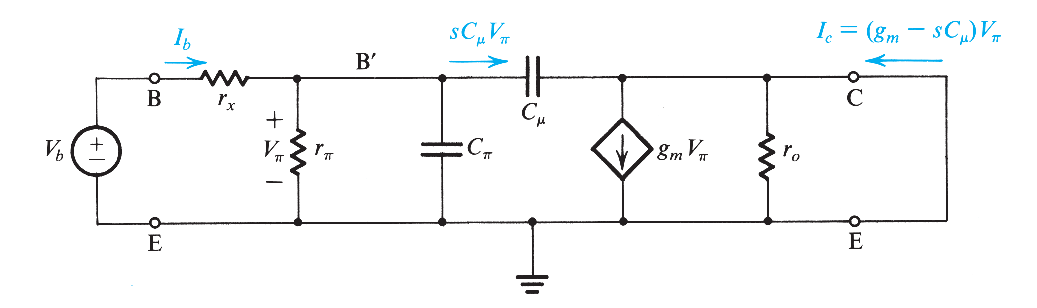


Figure 10.15 Circuit for deriving an expression for $h_{fe}(s) \equiv I_c/I_b$.

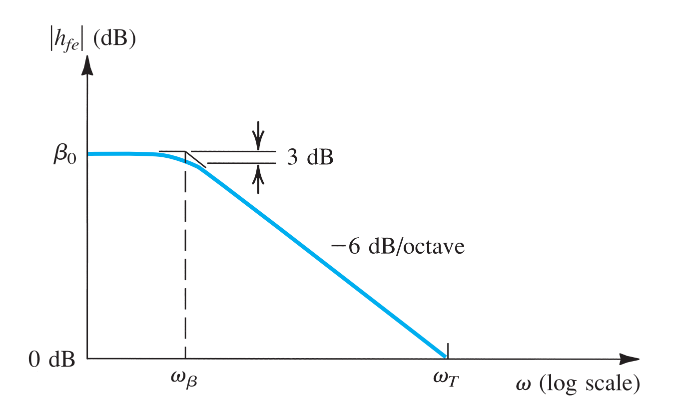


Figure 10.16 Bode plot for $|h_{fe}|$.

Thus,

$$h_{fe} = \frac{\beta_0}{1 + s(C_\pi + C_\mu)r_\pi} \tag{10.38}$$

where β_0 is the low-frequency value of β. Thus h_{fe} has a single-pole (or single-time-constant) response with a 3-dB frequency at $\omega = \omega_\beta$, where

$$\omega_\beta = \frac{1}{(C_\pi + C_\mu)r_\pi} \tag{10.39}$$

Figure 10.16 shows a Bode plot for $|h_{fe}|$. From the –6-dB/octave slope, it follows that the frequency at which $|h_{fe}|$ drops to unity, which is called the **unity-gain bandwidth** ω_T, is given by

$$\omega_T = \beta_0 \omega_\beta$$

Thus,

$$\omega_T = \frac{g_m}{C_\pi + C_\mu} \tag{10.40}$$

and

$$f_T = \frac{g_m}{2\pi(C_\pi + C_\mu)} \tag{10.41}$$

This expression is identical to that of f_T for the MOSFET (Eq. 10.29) with C_π replacing C_{gs} and C_μ replacing C_{gd}.

The unity-gain bandwidth f_T, also known as the **transition frequency**, which gives rise to the subscript T, is usually specified on the data sheets of a transistor. In some cases f_T is given as a function of I_C and V_{CE}. To see how f_T changes with I_C, recall that g_m is directly proportional to I_C, but only part of C_π (the diffusion capacitance C_{de}) is directly proportional to I_C. It follows that f_T decreases at low currents, as shown in Fig. 10.17. However, the decrease in f_T at high currents, also shown in Fig. 10.17, cannot be explained by this argument; rather, it is due to the same phenomenon that causes β_0 to decrease at high currents (Section 6.4.2). In the region where f_T is almost constant, C_π is dominated by the diffusion part and is much greater than C_μ. That is, $C_\pi + C_\mu \simeq C_{de} = \tau_F g_m$ and

$$f_T \simeq \frac{1}{2\pi \tau_F} \tag{10.42}$$

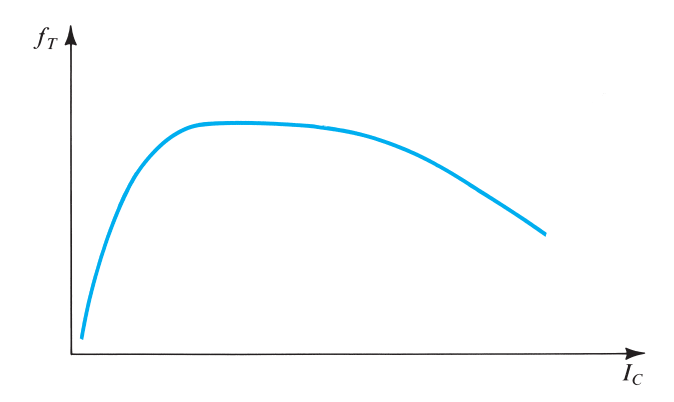


Figure 10.17 Variation of f_T with I_C.

Typically, f_T is in the range of 100 MHz to tens of gigahertz. The value of f_T can be used in Eq. (10.41) to determine $C_\pi + C_\mu$. The capacitance C_μ is usually determined separately by measuring the capacitance between base and collector at the desired reverse-bias voltage V_{CB}.

An important observation to make from the high-frequency model of Fig. 10.14(a) is that at frequencies above 5 to 10 f_β, one may ignore the resistance r_π. It can be seen then that r_x becomes the only resistive part of the input impedance at high frequencies. Thus r_x plays an important role in determining the frequency response of transistor circuits at high frequencies. It follows that an accurate determination of r_x can be made only from a high-frequency measurement.

Before leaving this section, we should mention that the high-frequency models of Fig. 10.14 characterize transistor operation fairly accurately up to a frequency of about $0.2 f_T$. At higher frequencies one has to add other parasitic elements to the model.

EXERCISES

10.5 Find C_{de}, C_{je}, C_π, C_μ, and f_T for a BJT operating at a dc collector current $I_C = 1$ mA and a CBJ reverse bias of 2 V. The device has $\tau_F = 20$ ps, $C_{je0} = 20$ fF, $C_{\mu 0} = 20$ fF, $V_{0e} = 0.9$ V, $V_{0c} = 0.5$ V, and $m_{CBJ} = 0.33$.
Ans. 0.8 pF; 40 fF; 0.84 pF; 12 fF; 7.47 GHz

10.6 For a BJT operated at $I_C = 1$ mA, determine f_T and C_π if $C_\mu = 2$ pF and $|h_{fe}| = 10$ at 50 MHz.
Ans. 500 MHz; 10.7 pF

10.7 If C_π of the BJT in Exercise 10.6 includes a relatively constant depletion-layer capacitance of 2 pF, find f_T of the BJT when operated at $I_C = 0.1$ mA.
Ans. 130.7 MHz

Summary For convenient reference, Table 10.2 provides a summary of the relationships used to determine the values of the parameters of the BJT high-frequency model.

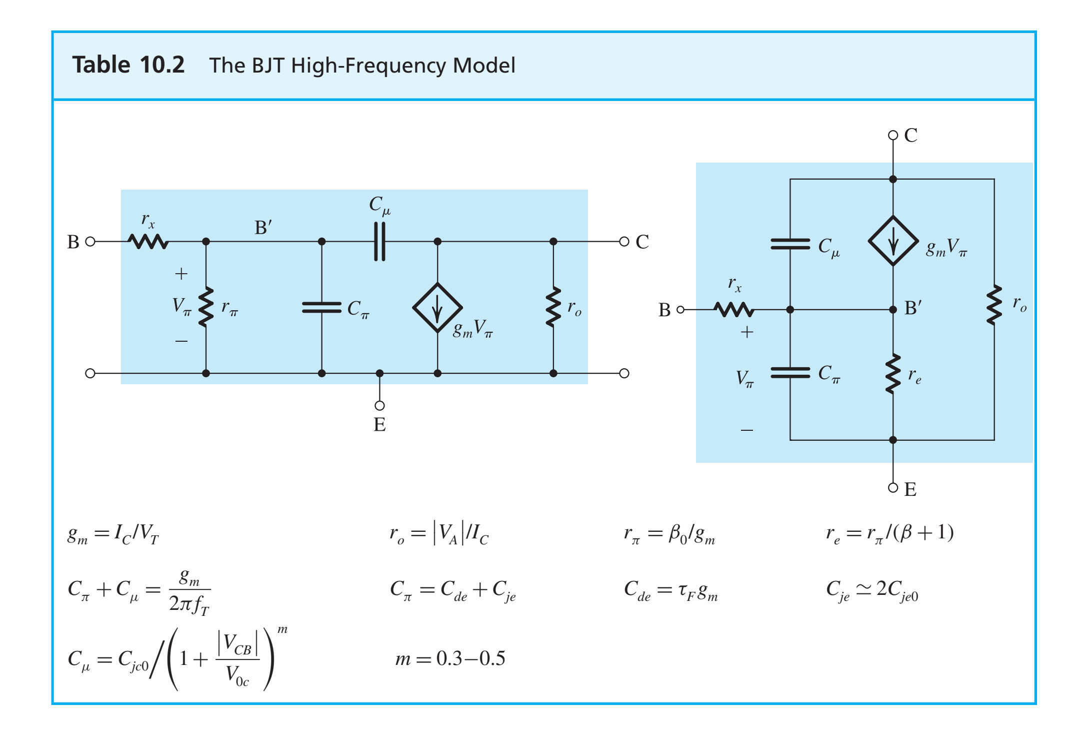


Table 10.2 The BJT High-Frequency Model

$g_m = I_C/V_T$ $\quad r_o = |V_A|/I_C$ $\quad r_\pi = \beta_0/g_m$ $\quad r_e = r_\pi/(\beta+1)$

$C_\pi + C_\mu = \dfrac{g_m}{2\pi f_T}$ $\quad C_\pi = C_{de} + C_{je}$ $\quad C_{de} = \tau_F g_m$ $\quad C_{je} \simeq 2C_{je0}$

$C_\mu = C_{jc0} \Big/ \left(1 + \dfrac{|V_{CB}|}{V_{0c}}\right)^m$ $\quad m = 0.3\text{–}0.5$

10.3 High-Frequency Response of the CS and CE Amplifiers

Equipped with equivalent-circuit models that represent the high-frequency operation of the MOSFET and the BJT, we now address the question of the high-frequency performance of the CS and CE amplifiers. Our objective is to identify the mechanism that limits the high-frequency performance of these important amplifier configurations. As well, we need to find a simple approach to estimate the frequency f_H at which the gain falls by 3 dB below its value at midband frequencies, $|A_M|$.

The analysis presented here applies equally well to discrete-circuit, capacitively coupled amplifiers and to IC amplifiers. The frequency response of the first was shown in Fig. 10.1 and that of the latter in Fig. 10.2. At the frequencies of interest to us here (the high-frequency band), all coupling and bypass capacitors behave as perfect short circuits, and amplifiers of both types have identical high-frequency equivalent circuits.

10.3.1 The Common-Source Amplifier

Figure 10.18(a) shows the high-frequency, equivalent-circuit model of a CS amplifier. It is obtained by replacing the MOSFET in an amplifier circuit such as that in Fig. 10.3(a) by its high-frequency, equivalent-circuit model of Fig. 10.6(c), while as always eliminating dc sources. Observe that the circuit in Fig. 10.18(a) is general; for instance, it includes a resistance R_G, which arises only in the case of a discrete-circuit amplifier such as that in Fig. 10.3(a), where $R_G \equiv R_{G1} \| R_{G2}$. Also, R_D can be either a passive resistance or the output resistance of a current-source load, and similarly for R_L.

The equivalent circuit of Fig. 10.18(a) can be simplified by utilizing Thévenin theorem at the input side and by combining the three parallel resistances at the output side. The resulting simplified circuit is shown in Fig. 10.18(b). The midband gain A_M can be found from this circuit by setting C_{gs} and C_{gd} to zero. The result is

$$A_M = \frac{V_o}{V_{\text{sig}}} = -\frac{R_G}{R_G + R_{\text{sig}}}(g_m R_L') \tag{10.43}$$

The circuit in Fig. 10.18(b) can be analyzed to obtain its transfer function V_o/V_{sig} in terms of the complex frequency variable s. Because two capacitors, C_{gs} and C_{gd}, are present, the resulting transfer function will be of second order. The poles and zeros can then be determined. This, however, will not provide us with simple expressions that reveal the essence of what limits the high-frequency operation of the CS amplifier. We need such insight in order to be able to make intelligent decisions when designing the circuit. Therefore, we shall not derive the transfer function and instead opt for an approximate approach that will reveal considerable information about the high-frequency operation of the CS amplifier.

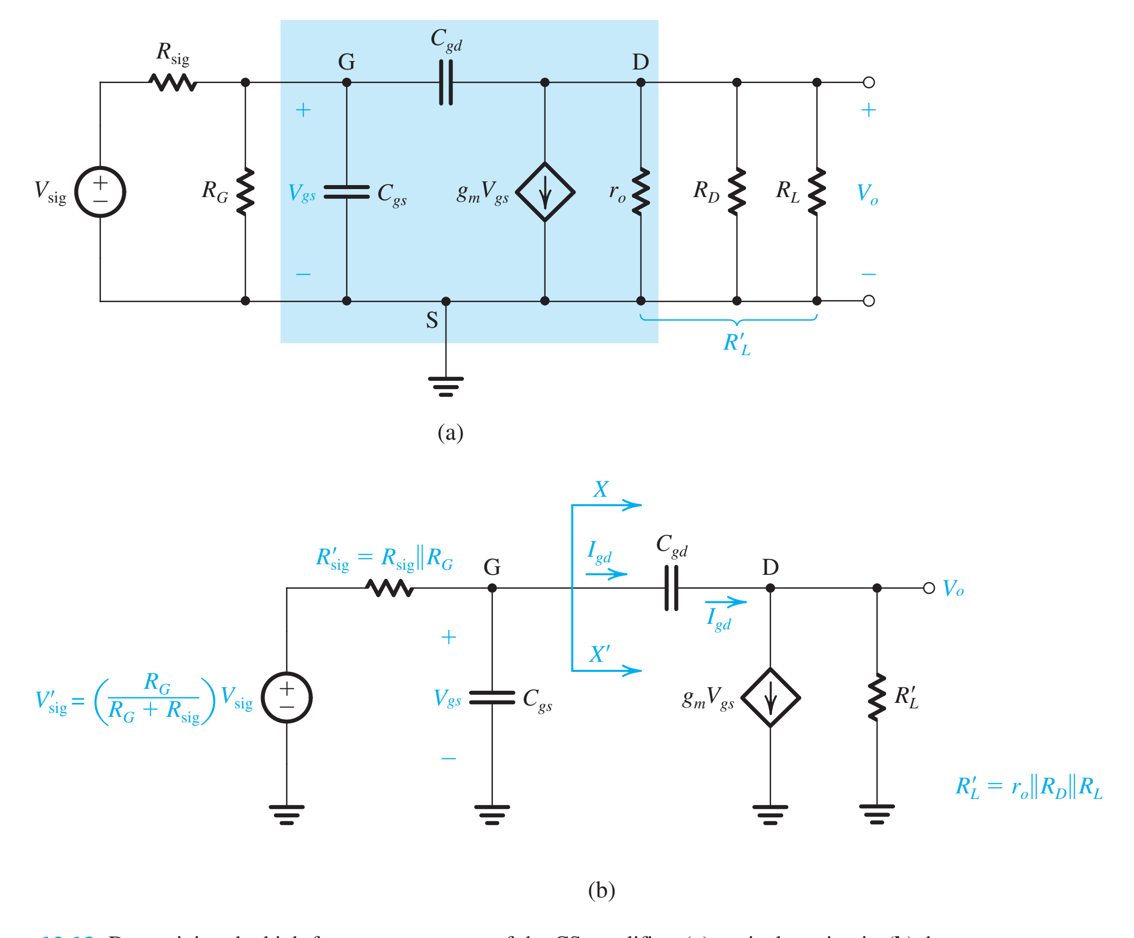
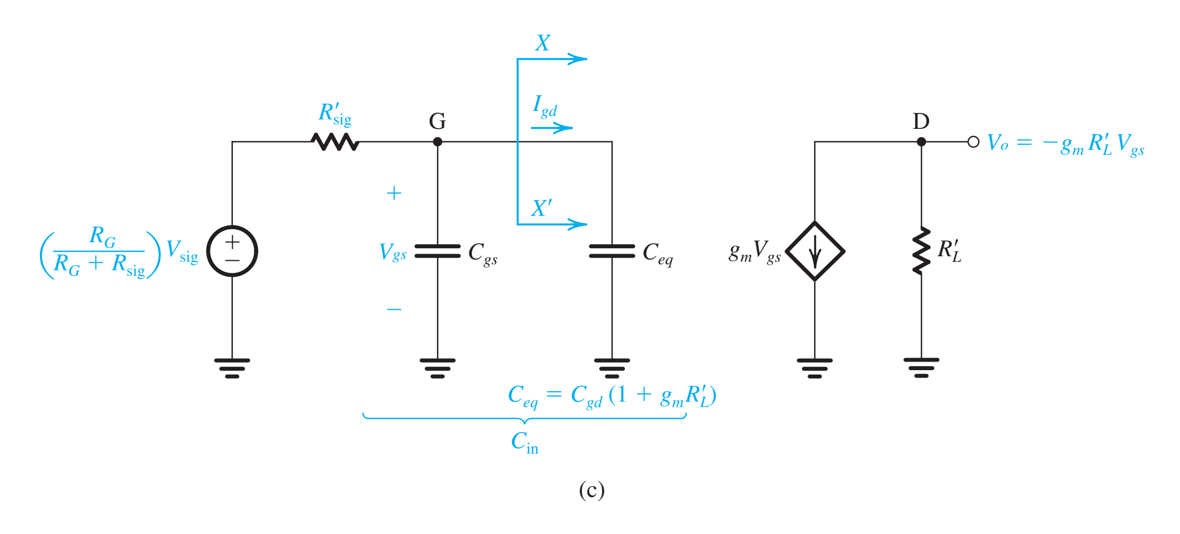
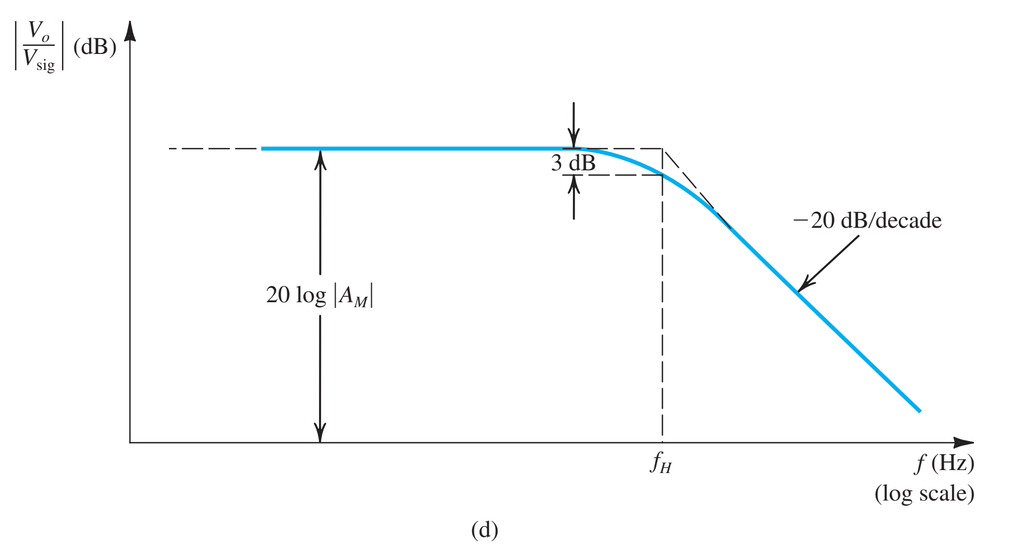


Figure 10.18 Determining the high-frequency response of the CS amplifier: (**a**) equivalent circuit; (**b**) the circuit of (**a**) simplified at the input and the output; (**c**) the equivalent circuit with C_{gd} replaced at the input side with the equivalent capacitance C_{eq}; (**d**) the frequency-response plot, which is that of a low-pass, single-time-constant circuit.



(c)



(d)

Figure 10.18 *continued*

Our approach is to focus on the input side of the circuit and seek to simplify the input circuit to a simple RC low-pass network. To do so we need to replace the bridging capacitor C_{gd} by an equivalent capacitance C_{eq} between node G and ground. Toward that end, consider first the output node. It can be seen that the load current is $(g_m V_{gs} - I_{gd})$, where $(g_m V_{gs})$ is the output current of the transistor and I_{gd} is the current supplied through the very small capacitance C_{gd}. At frequencies in the vicinity of f_H, which defines the edge of the midband, it is reasonable to assume that I_{gd} is still much smaller than $(g_m V_{gs})$, with the result that V_o can be given approximately by

$$V_o \simeq -(g_m V_{gs}) R'_L = -g_m R'_L V_{gs} \tag{10.44}$$

where

$$R'_L = r_o \| R_D \| R_L$$

Since $V_o = V_{ds}$, Eq. (10.44) indicates that the gain from gate to drain is $-g_m R_L'$, the same value as in the midband. The current I_{gd} can now be found as

$$\begin{aligned} I_{gd} &= sC_{gd}(V_{gs} - V_o) \\ &= sC_{gd}[V_{gs} - (-g_m R_L' V_{gs})] \\ &= sC_{gd}(1 + g_m R_L')V_{gs} \end{aligned}$$

Now, the left-hand side of the circuit in Fig. 10.18(b), at XX', knows of the existence of C_{gd} only through the current I_{gd}. Therefore, we can replace C_{gd} by an equivalent capacitance C_{eq} between the gate and ground as long as C_{eq} draws a current equal to I_{gd}. That is,

$$sC_{eq}V_{gs} = sC_{gd}(1 + g_m R_L')V_{gs}$$

which results in

$$C_{eq} = C_{gd}(1 + g_m R_L') \tag{10.45}$$

Thus C_{gd} gives rise to a much larger capacitance C_{eq}, which appears at the amplifier input. The multiplication effect that C_{gd} undergoes comes about because it is connected between circuit nodes G and D, whose voltages are related by a large negative gain ($-g_m R_L'$). This effect is known as the **Miller effect**, and $(1 + g_m R_L')$ is known as the **Miller multiplier**.

Using C_{eq} enables us to simplify the equivalent circuit at the input side to that shown in Fig. 10.18(c), which we recognize as a single-time-constant (STC) circuit of the low-pass type (Section 1.6 and Appendix E). Reference to Table 1.2 enables us to express the output voltage V_{gs} of the STC circuit in the form

$$V_{gs} = \left(\frac{R_G}{R_G + R_{\text{sig}}} V_{\text{sig}}\right) \frac{1}{1 + \dfrac{s}{\omega_0}} \tag{10.46}$$

where ω_0 is the corner frequency, the break frequency, or the pole frequency of the STC circuit,

$$\omega_0 = 1/C_{\text{in}} R_{\text{sig}}' \tag{10.47}$$

with

$$C_{\text{in}} = C_{gs} + C_{eq} = C_{gs} + C_{gd}(1 + g_m R_L') \tag{10.48}$$

and

$$R_{\text{sig}}' = R_{\text{sig}} \parallel R_G \tag{10.49}$$

Combining Eqs. (10.44) and (10.46) results in the following expression for the high-frequency gain of the CS amplifier,

$$\frac{V_o}{V_{\text{sig}}} = -\left(\frac{R_G}{R_G + R_{\text{sig}}}\right)(g_m R_L') \frac{1}{1 + \dfrac{s}{\omega_0}} \tag{10.50}$$

which can be expressed in the form

$$\frac{V_o}{V_{\text{sig}}} = \frac{A_M}{1 + \dfrac{s}{\omega_H}} \tag{10.51}$$

where the midband gain A_M is given by Eq. (10.43) and ω_H is the upper 3-dB frequency,

$$\omega_H = \omega_0 = \frac{1}{C_{in} R'_{sig}} \tag{10.52}$$

and



$$f_H = \frac{\omega_H}{2\pi} = \frac{1}{2\pi C_{in} R'_{sig}} \tag{10.53}$$

We thus see that the high-frequency response will be that of a low-pass STC network with a 3-dB frequency f_H determined by the time constant $C_{in}R'_{sig}$. Figure 10.18(d) shows a sketch of the magnitude of the high-frequency gain.

Before leaving this section we wish to make a number of observations:

1. The upper 3-dB frequency is determined by the interaction of $R'_{sig} = R_{sig} \| R_G$ and $C_{in} = C_{gs} + C_{gd}(1 + g_m R'_L)$. Since the bias resistance R_G is usually very large, it can be neglected, resulting in $R'_{sig} \simeq R_{sig}$, the resistance of the signal source. It follows that a large value of R_{sig} will cause f_H to be lowered.
2. The total input capacitance C_{in} is usually dominated by C_{eq}, which in turn is made large by the multiplication effect that C_{gd} undergoes. Thus, although C_{gd} is usually a very small capacitance, its effect on the amplifier frequency response can be very significant as a result of its multiplication by the factor $(1 + g_m R'_L)$, which is approximately equal to the midband gain of the amplifier. This is the Miller effect, which causes the CS amplifier to have a large total input capacitance C_{in} and hence a low f_H.
3. To extend the high-frequency response of a MOSFET amplifier, we have to find configurations in which the Miller effect is absent or at least reduced. We shall return to this subject at great length in Section 10.5 and beyond.
4. The above analysis, resulting in an STC or a single-pole response, is approximate. Specifically, it is based on neglecting I_{gd} relative to $g_m V_{gs}$, an assumption that applies well at frequencies not too much higher than f_H. An exact analysis of the circuit in Fig. 10.18(a) reveals that the circuit has a second pole with a frequency much greater than f_H, and transmission zeros at $s = \infty$ and $s = g_m/C_{gd}$; the latter's frequency is also much greater than f_H. Thus both the second pole and the zero will have negligible effect on our estimate of f_H. Thus, the method that uses the Miller effect is more than sufficient for a quick estimate of f_H. As well, the approximate approach helps to reveal the primary limitation on the high-frequency response: the Miller effect.
5. The CS amplifier is said to have a dominant high-frequency pole with frequency $f_P \simeq f_H$.

Example 10.3

Find the midband gain A_M and the upper 3-dB frequency f_H of a CS amplifier fed with a signal source having an internal resistance $R_{sig} = 100\ \text{k}\Omega$. The amplifier has $R_G = 4.7\ \text{M}\Omega$, $R_D = R_L = 15\ \text{k}\Omega$, $g_m = 1\ \text{mA/V}$, $r_o = 150\ \text{k}\Omega$, $C_{gs} = 1\ \text{pF}$, and $C_{gd} = 0.4\ \text{pF}$. Also, find the frequency of the transmission zero.

Solution

$$A_M = -\frac{R_G}{R_G + R_{sig}} g_m R_L'$$

where

$$R_L' = r_o \| R_D \| R_L = 150 \| 15 \| 15 = 7.14 \text{ k}\Omega$$
$$g_m R_L' = 1 \times 7.14 = 7.14 \text{ V/V}$$

Thus,

$$A_M = -\frac{4.7}{4.7 + 0.1} \times 7.14 = -7 \text{ V/V}$$

The equivalent capacitance, C_{eq}, is found as

$$C_{eq} = (1 + g_m R_L')C_{gd}$$
$$= (1 + 7.14) \times 0.4 = 3.26 \text{ pF}$$

The total input capacitance C_{in} can be now obtained as

$$C_{in} = C_{gs} + C_{eq} = 1 + 3.26 = 4.26 \text{ pF}$$

The upper 3-dB frequency f_H is found from

$$f_H = \frac{1}{2\pi C_{in}(R_{sig} \| R_G)}$$
$$= \frac{1}{2\pi \times 4.26 \times 10^{-12}(0.1 \| 4.7) \times 10^6}$$
$$= 382 \text{ kHz}$$

Finally, the transmission zero has a frequency

$$f_Z = \frac{g_m}{2\pi C_{gd}} = \frac{1 \times 10^{-3}}{2\pi \times 0.4 \times 10^{-12}} = 398 \text{ MHz}$$

which is more than 1000 times higher than f_H.

EXERCISES

10.8 For the CS amplifier specified in Example 10.3, find the values of A_M and f_H that result when the signal-source resistance is reduced to 10 kΩ.
Ans. –7.12 V/V; 3.7 MHz

10.9 If it is possible to replace the MOSFET used in the amplifier in Example 10.3 with another having the same C_{gs} but a smaller C_{gd}, what is the maximum value that its C_{gd} can be in order to obtain an f_H of at least 1 MHz?
Ans. 0.08 pF

10.3.2 The Common-Emitter Amplifier

Figure 10.19(a) shows the high-frequency equivalent circuit of a CE amplifier. It is obtained by replacing the BJT in a circuit such as that in Fig. 10.9(a) with its high-frequency, equivalent-circuit model of Fig. 10.14(a), and, as usual, eliminating all dc sources. Observe that the circuit in Fig. 10.19(a) is general and applies to both discrete- and integrated-circuit amplifiers. Thus, it includes R_B, which is usually present in discrete circuits. Also R_C can be either a passive resistance or the output resistance of a current-source load, and similarly for R_L.

The equivalent circuit of Fig. 10.19(a) can be simplified by utilizing Thévenin theorem at the input side and by combining the three parallel resistances at the output side. Specifically, the reader should be able to show that applying Thévenin theorem *twice* simplifies the resistive network at the input side to a signal generator V'_{sig} and a resistance R'_{sig}, with the values indicated in the figure.

The equivalent circuit in Fig. 10.19(b) can be used to obtain the midband gain A_M by setting C_π and C_μ to zero. The result is

$$A_M = \frac{V_o}{V_{sig}} = -\frac{R_B}{R_B + R_{sig}} \frac{r_\pi}{r_\pi + r_x + (R_{sig} \| R_B)} (g_m R'_L) \tag{10.54}$$

where

$$R'_L = r_o \| R_C \| R_L \tag{10.55}$$

Next we observe that the circuit in Fig. 10.19(b) is identical to that of the CS amplifier in Fig. 10.18(b). Thus the analysis can follow the same process we used for the CS case. The analysis is illustrated in Fig. 10.19(c) and (d). The final result is that the CE amplifier gain at high frequencies is given approximately by

$$\frac{V_o}{V_{sig}} = \frac{A_M}{1 + \frac{s}{\omega_H}} \tag{10.56}$$

where A_M is given by Eq. (10.54) and the 3-dB frequency f_H is given by

$$f_H = \frac{\omega_H}{2\pi} = \frac{1}{2\pi C_{in} R'_{sig}} \tag{10.57}$$

where

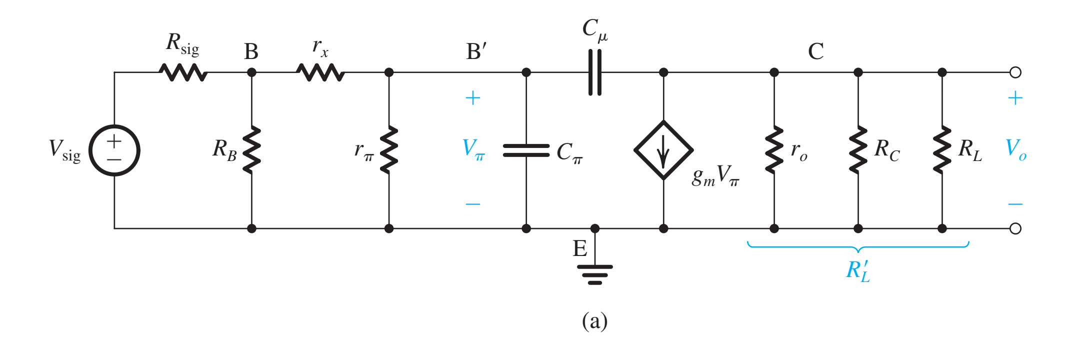
$$C_{in} = C_\pi + C_\mu (1 + g_m R'_L) \tag{10.58}$$

and

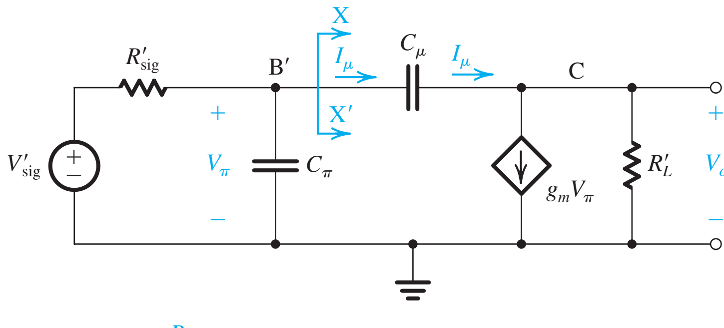
$$R'_{sig} = r_\pi \| [r_x + (R_B \| R_{sig})] \tag{10.59}$$

Observe that C_{in} is simply the sum of C_π and the Miller capacitance $C_\mu(1 + g_m R'_L)$. The resistance R'_{sig} seen by C_{in} can be easily found from the circuit in Fig. 10.19(a) as follows: Reduce V_{sig} to zero, "grab hold" of the terminals B′ and E and look back (to the left). You will see r_π in parallel with r_x, which is in series with $(R_B \| R_{sig})$. This way of finding the resistance "seen by a capacitance" is very useful and spares one from tedious work!

Finally, comments very similar to those made on the high-frequency response of the CS amplifier can be made here as well.



(a)

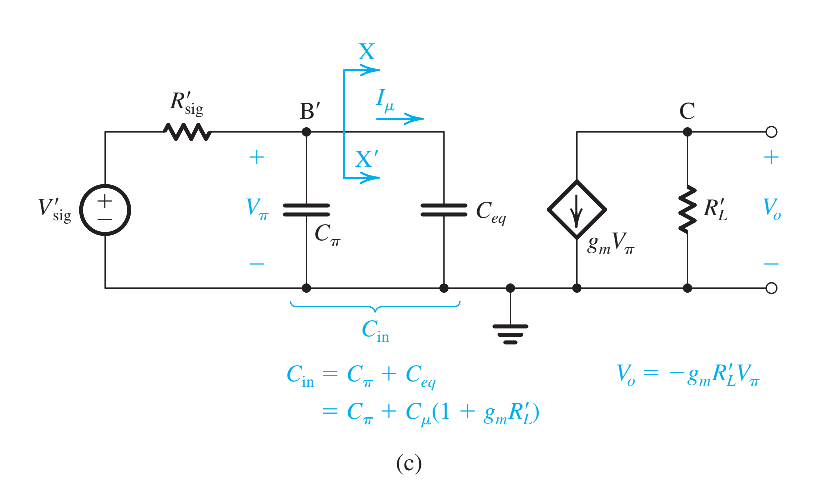


$$V'_{sig} = V_{sig}\frac{R_B}{R_B + R_{sig}}\,\frac{r_\pi}{r_\pi + r_x + (R_{sig}\|R_B)}$$

$$R'_L = r_o\|R_C\|R_L$$

$$R'_{sig} = r_\pi\|[r_x + (R_B\|R_{sig})]$$

(b)



$$C_{in} = C_\pi + C_{eq}$$
$$= C_\pi + C_\mu(1 + g_m R'_L)$$

$$V_o = -g_m R'_L V_\pi$$

(c)

Figure 10.19 Determining the high-frequency response of the CE amplifier: (**a**) equivalent circuit; (**b**) the circuit of (**a**) simplified at both the input side and the output side; (**c**) equivalent circuit with C_μ replaced at the input side with the equivalent capacitance C_{eq}; (**d**) sketch of the frequency-response plot, which is that of a low-pass STC circuit.

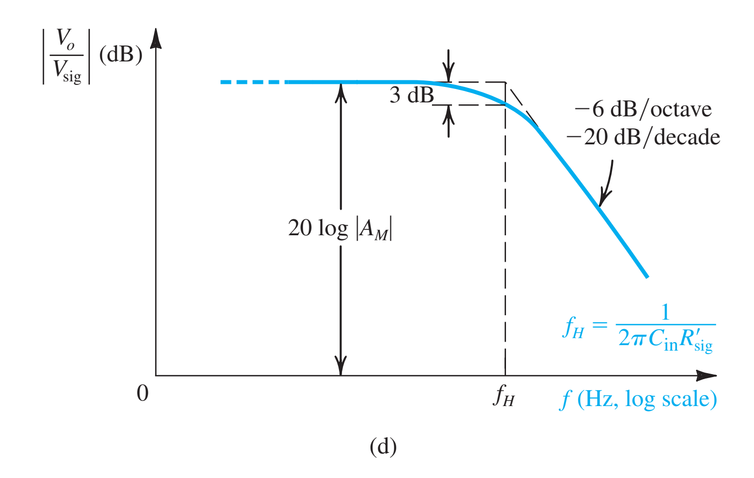


Figure 10.19 *continued*

Example 10.4

It is required to find the midband gain and the upper 3-dB frequency of the common-emitter amplifier of Fig. 10.9(a) for the following case: $I_E = 1$ mA, $R_B = R_{B1} \| R_{B2} = 100$ kΩ, $R_C = 8$ kΩ, $R_{sig} = 5$ kΩ, $R_L = 5$ kΩ, $\beta_0 = 100$, $V_A = 100$ V, $C_\mu = 1$ pF, $f_T = 800$ MHz, and $r_x = 50$ Ω. Also, determine the frequency of the transmission zero.

Solution

The transistor is biased at $I_C \simeq 1$ mA. Thus the values of its hybrid-π model parameters are

$$g_m = \frac{I_C}{V_T} = \frac{1 \text{ mA}}{25 \text{ mV}} = 40 \text{ mA/V}$$

$$r_\pi = \frac{\beta_0}{g_m} = \frac{100}{40 \text{ mA/V}} = 2.5 \text{ k}\Omega$$

$$r_o = \frac{V_A}{I_C} = \frac{100 \text{ V}}{1 \text{ mA}} = 100 \text{ k}\Omega$$

$$C_\pi + C_\mu = \frac{g_m}{\omega_T} = \frac{40 \times 10^{-3}}{2\pi \times 800 \times 10^6} = 8 \text{ pF}$$

$$C_\mu = 1 \text{ pF}$$

$$C_\pi = 7 \text{ pF}$$

$$r_x = 50 \ \Omega$$

The midband voltage gain is

$$A_M = -\frac{R_B}{R_B + R_{sig}} \frac{r_\pi}{r_\pi + r_x + (R_B \| R_{sig})} g_m R'_L$$

where

$$R'_L = r_o \| R_C \| R_L$$
$$= (100 \| 8 \| 5)\text{k}\Omega = 3\text{ k}\Omega$$

Thus,

$$g_m R'_L = 40 \times 3 = 120\text{ V/V}$$

and

$$A_M = -\frac{100}{100+5} \times \frac{2.5}{2.5 + 0.05 + (100 \| 5)} \times 120$$
$$= -39\text{ V/V}$$

and

$$20\log|A_M| = 32\text{ dB}$$

To determine f_H we first find C_{in},

$$C_{in} = C_\pi + C_\mu(1 + g_m R'_L)$$
$$= 7 + 1(1 + 120) = 128\text{ pF}$$

and the effective source resistance R'_{sig},

$$R'_{sig} = r_\pi \| \left[r_x + \left(R_B \| R_{sig}\right)\right]$$
$$= 2.5 \| [0.05 + (100 \| 5)]$$
$$= 1.65\text{ k}\Omega$$

Thus,

$$f_H = \frac{1}{2\pi C_{in} R'_{sig}} = \frac{1}{2\pi \times 128 \times 10^{-12} \times 1.65 \times 10^3} = 754\text{ kHz}$$

Finally, as in the case of the CS amplifier, it can be shown that the CE amplifier has a transmission zero with frequency

$$f_Z = \frac{g_m}{2\pi C_\mu} = \frac{40 \times 10^{-3}}{2\pi \times 1 \times 10^{-12}} = 6.37\text{ GHz}$$

which is much higher than f_H.

EXERCISE

10.10 For the amplifier in Example 10.4, find the value of R_L that reduces the midband gain to half the value found. What value of f_H results? Note the trade-off between gain and bandwidth.
Ans. 1.9 kΩ; 1.42 MHz

10.3.3 Miller's Theorem

In our analysis of the high-frequency response of the common-source and common-emitter amplifiers, we employed a technique for replacing the bridging capacitance (C_{gs} or C_μ) by an equivalent input capacitance. This very useful and effective technique is based on a general theorem known as **Miller's theorem**, which we now present.

Consider the situation in Fig. 10.20(a). As part of a larger circuit that is not shown, we have isolated two circuit nodes, labeled 1 and 2, between which an impedance Z is connected. Nodes 1 and 2 are also connected to other parts of the circuit, as signified by the broken lines emanating from the two nodes. Furthermore, it is assumed that somehow it has been determined that the voltage at node 2 is related to that at node 1 by

$$V_2 = KV_1 \tag{10.60}$$

In typical situations, K is a gain factor that can be positive or negative and has a magnitude usually larger than unity. This, however, is not an assumption for Miller's theorem.

Miller's theorem states that impedance Z can be replaced by two impedances: Z_1 connected between node 1 and ground and Z_2 connected between node 2 and ground, where

$$Z_1 = Z/(1-K) \tag{10.61a}$$

and

$$Z_2 = Z\bigg/\left(1-\frac{1}{K}\right) \tag{10.61b}$$

to obtain the equivalent circuit shown in Fig. 10.20(b).

The proof of Miller's theorem is achieved by deriving Eqs. (10.61) as follows: In the original circuit of Fig. 10.20(a), the only way that node 1 "feels the existence" of impedance Z is through the current I that Z draws away from node 1. Therefore, to keep this current unchanged in the equivalent circuit, we must choose the value of Z_1 so that it draws an equal current,

$$I_1 = \frac{V_1}{Z_1} = I = \frac{V_1 - KV_1}{Z}$$

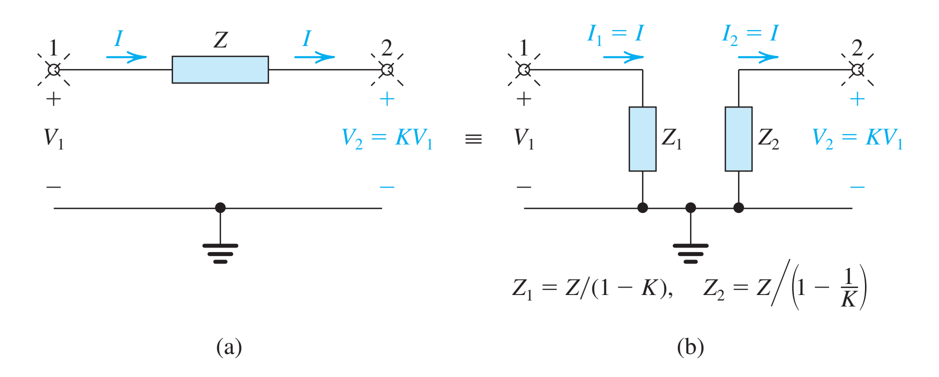


Figure 10.20 The Miller equivalent circuit.

which yields the value of Z_1 in Eq. (10.61a). Similarly, to keep the current into node 2 unchanged, we must choose the value of Z_2 so that

$$I_2 = \frac{0 - V_2}{Z_2} = \frac{0 - KV_1}{Z_2} = I = \frac{V_1 - KV_1}{Z}$$

which yields the expression[5] for Z_2 in Eq. (10.61b).

Example 10.5

Figure 10.21(a) shows an ideal voltage amplifier having a gain of −100 V/V with an impedance Z connected between its output and input terminals. Find the Miller equivalent circuit when Z is (a) a 1-MΩ resistance and (b) a 1-pF capacitance. In each case, use the equivalent circuit to determine V_o/V_{sig}.

Solution

(a) For $Z = 1$ MΩ, employing Miller's theorem results in the equivalent circuit in Fig. 10.21(b), where

$$Z_1 = \frac{Z}{1 - K} = \frac{1000 \text{ k}\Omega}{1 + 100} = 9.9 \text{ k}\Omega$$

$$Z_2 = \frac{Z}{1 - \dfrac{1}{K}} = \frac{1 \text{ M}\Omega}{1 + \dfrac{1}{100}} = 0.99 \text{ M}\Omega$$

The voltage gain can be found as follows:

$$\frac{V_o}{V_{sig}} = \frac{V_o}{V_i}\frac{V_i}{V_{sig}} = -100 \times \frac{Z_1}{Z_1 + R_{sig}}$$

$$= -100 \times \frac{9.9}{9.9 + 10} = -49.7 \text{ V/V}$$

(b) For Z as a 1-pF capacitance—that is, $Z = 1/sC = 1/s \times 1 \times 10^{-12}$ —applying Miller's theorem allows us to replace Z by Z_1 and Z_2, where

$$Z_1 = \frac{Z}{1 - K} = \frac{1/sC}{1 + 100} = 1/s(101C)$$

$$Z_2 = \frac{Z}{1 - \dfrac{1}{K}} = \frac{1}{1.01}\frac{1}{sC} = \frac{1}{s(1.01C)}$$

[5]Although not highlighted, the Miller equivalent circuit derived above is valid only as long as the rest of the circuit remains unchanged; otherwise the ratio of V_2 to V_1 might change. It follows that the Miller equivalent circuit *cannot* be used directly to determine the output resistance of an amplifier. This is because in determining output resistances it is implicitly assumed that the source signal is reduced to zero and that a test-signal source (voltage or current) is applied to the output terminals—obviously a major change in the circuit, rendering the Miller equivalent circuit no longer valid.

Example 10.5 *continued*

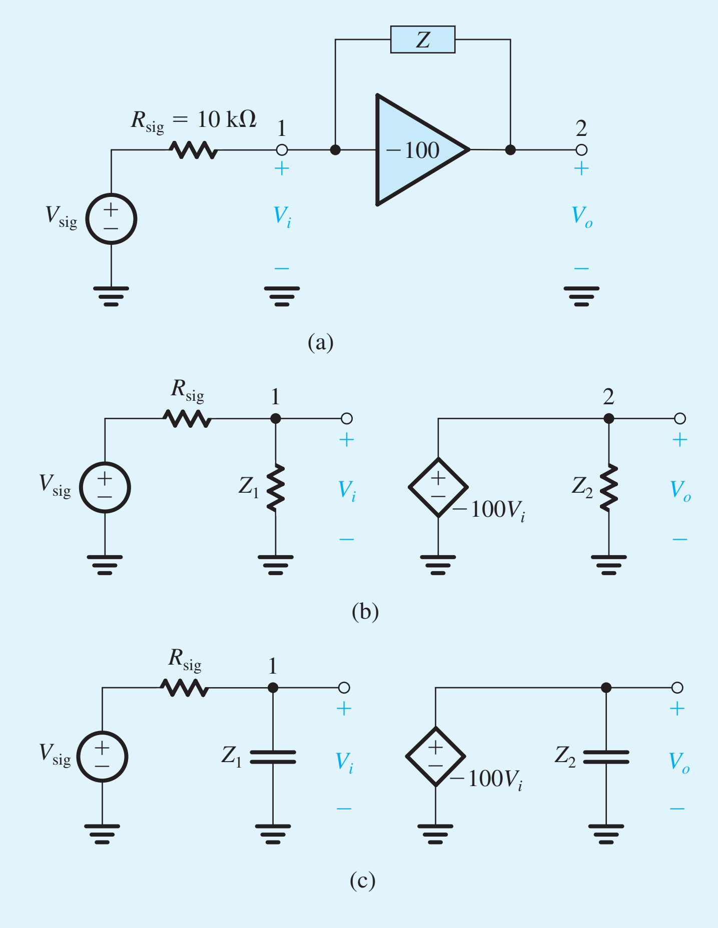


Figure 10.21 Circuits for Example 10.5.

It follows that Z_1 is a capacitance $101C = 101$ pF and that Z_2 is a capacitance $1.01C = 1.01$ pF. The resulting equivalent circuit is shown in Fig. 10.21(c), from which the voltage gain can be found as follows:

$$\frac{V_o}{V_{sig}} = \frac{V_o}{V_i}\frac{V_i}{V_{sig}} = -100\frac{1/sC_1}{1/(sC_1) + R_{sig}}$$

$$= \frac{-100}{1 + sC_1R_{sig}}$$

$$= \frac{-100}{1 + s \times 101 \times 1 \times 10^{-12} \times 10 \times 10^3}$$

$$= \frac{-100}{1 + s \times 1.01 \times 10^{-6}}$$

This is the transfer function of a first-order low-pass network with a dc gain of –100 and a 3-dB frequency f_{3dB} of

$$f_{3dB} = \frac{1}{2\pi \times 1.01 \times 10^{-6}} = 157.6 \text{ kHz}$$

From Example 10.5, we observe that the Miller replacement of a feedback or bridging resistance results, for a negative K, in a smaller resistance [by a factor $(1-K)$] at the input. If the feedback element is a capacitance, its value is multiplied by $(1-K)$ to obtain the equivalent capacitance at the input side. The multiplication of a feedback capacitance by $(1-K)$ is referred to as **Miller multiplication** or **Miller effect**. We have encountered the Miller effect in the analysis of the CS and CE amplifiers. Note, however, that we neglected the Miller capacitance at the output because it is small; for the CS case, $C_2 = C_{gd}\left(1+\frac{1}{g_m R'_L}\right) \simeq C_{gd}$.

JOHN MILTON MILLER—CAPACITANCE MULTIPLICATION:

In 1920 in the Scientific Papers of the National Bureau of Standards, John Miller first published data and analysis on the input capacitance of a vacuum-tube triode amplifier and the phenomenon ultimately known as the Miller effect. Attempts to eliminate or reduce this effect in amplifiers and oscillators led to several developments: First, in 1926, came vacuum-tube pentodes with internal grounded shielding elements; much later, in 1939, the cascode configuration was introduced, initially using vacuum-tube triodes and then with BJTs and MOSFETs.

10.3.4 Frequency Response of the CS Amplifier When R_{sig} Is Low

There are applications in which the CS amplifier is fed with a low-resistance signal source. In such a case, the high-frequency gain will no longer be limited by the interaction of the source resistance and the input capacitance. Rather, the high-frequency limitation occurs at the amplifier output as we shall now show.

Figure 10.22(a) shows the high-frequency equivalent circuit of the common-source amplifier in the limiting case when R_{sig} is zero. Observe that we have included a capacitance C_L across the load R_L. We have done this in the anticipation that a capacitance at the output node, even if it is very small, will play an important role in this case. Also, there always is some capacitance between the output node and ground. This can include C_{db} of the MOSFET [see the MOSFET high-frequency model of Fig. 10.12(b)], the input capacitance of another amplifier stage our amplifier feeds, other stray capacitances, and so on. Finally, we note that we did not include C_L in the previous analysis because its role is not significant when R_{sig} is large.

Returning to the circuit in Fig. 10.22(a), we can now derive its transfer function V_o/V_{sig}. First, note that

$$V_{gs} = V_{sig} \tag{10.62}$$

Second, the current I_{gd} that flows through C_{gd} can be expressed as

$$I_{gd} = sC_{gd}(V_{gs} - V_o) \tag{10.63}$$

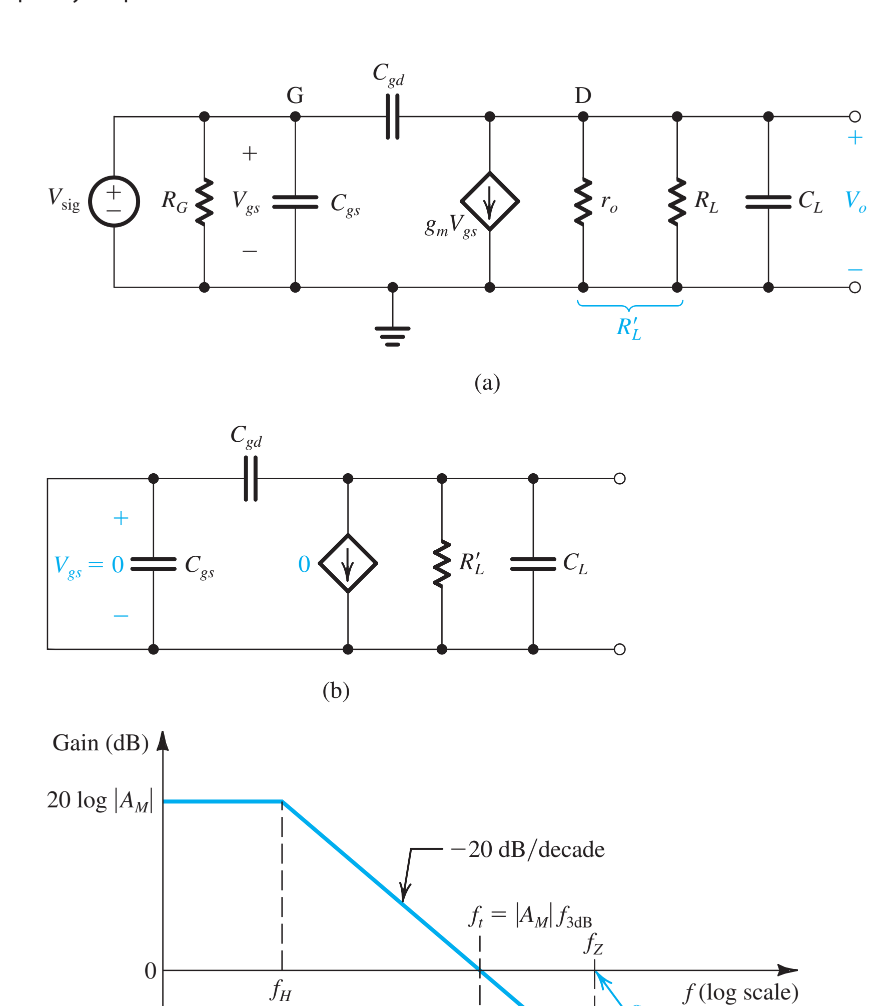


Figure 10.22 **(a)** High-frequency equivalent circuit of a CS amplifier fed with a signal source having a very low (effectively zero) resistance. **(b)** The circuit with V_{sig} reduced to zero. **(c)** Bode plot for the gain of the circuit in **(a)**.

Next, we can write a node equation at the output node as

$$I_{gd} = g_m V_{gs} + \frac{V_o}{R'_L} + sC_L V_o \tag{10.64}$$

where

$$R'_L = R_L \,\|\, r_o$$

Combining Eqs. (10.63) and (10.64) to eliminate I_{gd}, substituting $V_{gs} = V_{sig}$ from Eq. (10.62), and manipulating the resulting equation to obtain the transfer function V_o/V_{sig} results in

$$\frac{V_o}{V_{sig}} = -g_m R'_L \frac{1 - s(C_{gd}/g_m)}{1 + s(C_L + C_{gd})R'_L} \tag{10.65}$$

Thus, while the dc gain from gate to drain remains equal to $g_m R'_L$, and the frequency of the transmission zero remains unchanged at

$$f_Z = \frac{g_m}{2\pi C_{gd}} \tag{10.66}$$

the high-frequency response is now determined by a pole formed by $(C_L + C_{gd})$ together with R'_L. Thus the 3-dB frequency f_H is now given by

$$f_H = \frac{1}{2\pi (C_L + C_{gd}) R'_L} \tag{10.67}$$

To see how this pole is formed, refer to Fig. 10.22(b), which shows the equivalent circuit with V_{sig} reduced to zero. Observe that the circuit reduces to a capacitance $(C_L + C_{gd})$ in parallel with a resistance R'_L.

We note that the transmission zero frequency, given by Eq. (10.66), is much higher than f_H,

$$\frac{f_Z}{f_H} = (g_m R'_L)\left(1 + \frac{C_L}{C_{gd}}\right) \tag{10.68}$$

Thus, f_Z does not play a significant role in the vicinity of f_H. In fact, the gain decreases from its low-frequency value of $(g_m R'_L)$ at a uniform rate of –6 dB/octave (–20 dB/decade), reaching unity (0 dB) at a frequency f_t, which is equal to the **gain–bandwidth product**,

$$f_t = |A_M| f_H$$
$$= g_m R'_L \frac{1}{2\pi (C_L + C_{gd}) R'_L}$$

Thus,

$$f_t = \frac{g_m}{2\pi (C_L + C_{gd})} \tag{10.69}$$

Figure 10.22(c) shows a sketch of the high-frequency gain of the CS amplifier.

Example 10.6

Consider an IC CS amplifier fed with a source having $R_{sig} = 0$ and having an effective load resistance R'_L composed of r_o of the amplifier transistor in parallel with an equal resistance r_o of the current-source load. Let $g_m = 1.25$ mA/V, $r_o = 20\,\text{k}\Omega$, $C_{gs} = 20$ fF, $C_{gd} = 5$ fF, and $C_L = 25$ fF. Find A_M, f_H, f_t, and f_Z. If the amplifying transistor is to be operated at twice the original overdrive voltage while W and L remain unchanged, by what factor must the bias current be changed? What are the new values of A_M, f_H, f_t, and f_Z?

Solution

The low-frequency gain A_M is given by

$$A_M = -g_m R'_L = -g_m (r_o \| r_o)$$

Example 10.6 *continued*

Thus,

$$A_M = -\frac{1}{2} g_m r_o = -\frac{1}{2} \times 1.25 \times 20$$
$$= -12.5 \text{ V/V}$$

The 3-dB frequency f_H can be found using Eq. (10.67),

$$f_H = \frac{1}{2\pi(C_L + C_{gd})R'_L}$$
$$= \frac{1}{2\pi(25+5) \times 10^{-15} \times 10 \times 10^3}$$
$$= 530.5 \text{ MHz}$$

and the unity-gain frequency, which is equal to the gain–bandwidth product, can be determined as

$$f_t = |A_M| f_H = 12.5 \times 530.5 = 6.63 \text{ GHz}$$

The frequency of the zero is obtained using Eq. (10.66) as

$$f_Z = \frac{1}{2\pi} \frac{g_m}{C_{gd}}$$

$$= \frac{1}{2\pi} \frac{1.25 \times 10^{-3}}{5 \times 10^{-15}} \simeq 40 \text{ GHz}$$

Now, to double V_{OV}, I_D must be quadrupled. The new values of g_m and R'_L can be found as follows:

$$g_m = \frac{I_D}{V_{OV}/2} = 2.5 \text{ mA/V}$$

$$R'_L = \frac{1}{4} \times 10 = 2.5 \text{ k}\Omega$$

Thus the new value of A_M becomes

$$A_M = -g_m R'_L = -2.5 \times 2.5 = -6.25 \text{ V/V}$$

That of f_H becomes

$$f_H = \frac{1}{2\pi(C_L + C_{gd})R'_L}$$
$$= \frac{1}{2\pi(25+5) \times 10^{-15} \times 2.5 \times 10^3}$$
$$= 2.12 \text{ GHz}$$

and the unity-gain frequency (i.e., the gain–bandwidth product) becomes

$$f_t = 6.25 \times 2.12 = 13.3 \text{ GHz}$$

We note that doubling V_{OV} results in reducing the dc gain by a factor of 2 and increasing the bandwidth by a factor of 4. Thus, the gain–bandwidth product is doubled—a good bargain!

Finally, the frequency of the transmission zero f_Z will be doubled, becoming 80 GHz.

EXERCISE

10.11 For the CS amplifier considered in Example 10.6 operating at the original values of V_{OV} and I_D, find the value to which C_L should be increased to place f_t at 2 GHz.
Ans. 94.5 fF

We conclude this section by noting that a CE amplifier fed with $R_{sig} = 0$ can be analyzed in exactly the same way used for the CS case. The most general case when R_{sig} is not zero and C_L is present will be dealt with in the next section.

10.4 Useful Tools for the Analysis of the High-Frequency Response of Amplifiers

In Section 10.3 we presented an approximate method, utilizing the Miller effect, to analyze the high-frequency response of the CS and CE amplifiers; the method provides a reasonably accurate estimate of f_H and, equally important, considerable insight into the mechanism that limits high-frequency operation. Unfortunately, however, this method cannot deal with the case when a load capacitance C_L is present. As well, the method is not easily extendable to more complex amplifier circuits. For this reason, we will digress briefly into how to equip ourselves with a number of tools that will prove useful in the analysis of more complex circuits such as the cascode amplifier. We will begin by stepping back and more generally considering the amplifier high-frequency transfer function.

10.4.1 The High-Frequency Gain Function

The amplifier gain, taking into account the internal transistor capacitances, can be expressed as a function of the complex-frequency variable s in the general form

$$A(s) = A_M F_H(s) \tag{10.70}$$

where A_M is the midband gain, which for IC amplifiers is also equal to the low-frequency or dc gain (refer to Fig. 10.2). The value of A_M can be determined by analyzing the amplifier equivalent circuit while neglecting the effect of the transistor internal capacitances—that is, by assuming that they act as perfect open circuits. By taking these capacitances into account, we see that the gain acquires the factor $F_H(s)$, which can be expressed in terms of its poles and zeros, which are usually real, as follows:

$$F_H(s) = \frac{(1 + s/\omega_{Z1})(1 + s/\omega_{Z2})\cdots(1 + s/\omega_{Zn})}{(1 + s/\omega_{P1})(1 + s/\omega_{P2})\cdots(1 + s/\omega_{Pn})} \tag{10.71}$$

where $\omega_{P1}, \omega_{P2}, \ldots, \omega_{Pn}$ are positive numbers representing the frequencies of the n real poles[6] and $\omega_{Z1}, \omega_{Z2}, \ldots, \omega_{Zn}$ are positive, negative, or infinite numbers representing the frequencies of the n real transmission zeros. Since the frequencies of the zeros and poles are by definition greater than the midband frequencies, we see from Eq. (10.71) that as s approaches midband frequencies, $F_H(s)$ approaches unity and the gain approaches A_M.

10.4.2 Determining the 3-dB Frequency f_H

The amplifier designer usually is particularly interested in the part of the high-frequency band that is close to the midband. This is because the designer needs to estimate—and, if need be, modify—the value of the upper 3-dB frequency f_H (or ω_H; $f_H = \omega_H/2\pi$). Toward that end, it should be mentioned that in many cases the zeros are either at infinity or such high frequencies as to be of little significance to the determination of ω_H. If in addition one of the poles, say ω_{P1}, is of much lower frequency than any of the other poles, then this pole will have the greatest effect on the value of the amplifier ω_H. In other words, this pole will *dominate* the high-frequency response of the amplifier, and the amplifier is said to have a **dominant-pole response**. In such cases, the function $F_H(s)$ can be approximated by

$$F_H(s) \simeq \frac{1}{1 + s/\omega_{P1}} \tag{10.72}$$

which is the transfer function of a first-order (or STC) low-pass network (Appendix E). It follows that if a dominant pole exists, then the determination of ω_H is greatly simplified;

$$\omega_H \simeq \omega_{P1} \tag{10.73}$$

This is the situation we encountered in the cases of the common-source and common-emitter amplifiers analyzed in Section 10.3. As a rule of thumb, *a dominant pole exists if the lowest-frequency pole is at least two octaves (a factor of 4) away from the nearest pole or zero.*

If a dominant pole does not exist, the 3-dB frequency ω_H can be determined from a plot of $|F_H(j\omega)|$. Alternatively, an approximate formula for ω_H can be derived as follows. Consider, for simplicity, the case of a circuit having two poles and two zeros in the high-frequency band; that is,

$$F_H(s) = \frac{(1 + s/\omega_{Z1})(1 + s/\omega_{Z2})}{(1 + s/\omega_{P1})(1 + s/\omega_{P2})} \tag{10.74}$$

Substituting $s = j\omega$ and taking the squared magnitude gives

$$|F_H(j\omega)|^2 = \frac{\left(1 + \omega^2/\omega_{Z1}^2\right)\left(1 + \omega^2/\omega_{Z2}^2\right)}{\left(1 + \omega^2/\omega_{P1}^2\right)\left(1 + \omega^2/\omega_{P2}^2\right)}$$

[6]In certain cases, some of the poles can be complex. One notable such case is the source and emitter followers, which can have a pair of complex-conjugate poles (Section 10.6).

By definition, at $\omega = \omega_H$, $|F_H|^2 = \frac{1}{2}$; thus,

$$\frac{1}{2} = \frac{\left(1+\omega_H^2/\omega_{Z1}^2\right)\left(1+\omega_H^2/\omega_{Z2}^2\right)}{\left(1+\omega_H^2/\omega_{P1}^2\right)\left(1+\omega_H^2/\omega_{P2}^2\right)}$$

$$= \frac{1+\omega_H^2\left(\dfrac{1}{\omega_{Z1}^2}+\dfrac{1}{\omega_{Z2}^2}\right)+\omega_H^4/\omega_{Z1}^2\omega_{Z2}^2}{1+\omega_H^2\left(\dfrac{1}{\omega_{P1}^2}+\dfrac{1}{\omega_{P2}^2}\right)+\omega_H^4/\omega_{P1}^2\omega_{P2}^2} \tag{10.75}$$

Since ω_H is usually smaller than the frequencies of all the poles and zeros, we may neglect the terms containing ω_H^4 and solve for ω_H to obtain

$$\omega_H \simeq 1\Bigg/\sqrt{\frac{1}{\omega_{P1}^2}+\frac{1}{\omega_{P2}^2}-\frac{2}{\omega_{Z1}^2}-\frac{2}{\omega_{Z2}^2}} \tag{10.76}$$

This relationship can be extended to any number of poles and zeros as

$$\omega_H \simeq 1\Bigg/\sqrt{\left(\frac{1}{\omega_{P1}^2}+\frac{1}{\omega_{P2}^2}+\cdots\right)-2\left(\frac{1}{\omega_{Z1}^2}+\frac{1}{\omega_{Z2}^2}+\cdots\right)} \tag{10.77}$$

Note that if one of the poles, say P_1, is dominant, then $\omega_{P1} \ll \omega_{P2}, \omega_{P3}, \ldots, \omega_{Z1}, \omega_{Z2}, \ldots$, and Eq. (10.77) reduces to Eq. (10.73).

Example 10.7

The high-frequency response of an amplifier is characterized by the transfer function

$$F_H(s) = \frac{1-s/10^5}{\left(1+s/10^4\right)\left(1+s/4\times 10^4\right)}$$

Determine the 3-dB frequency approximately and exactly.

Solution

Noting that the lowest-frequency pole at 10^4 rad/s is two octaves lower than the second pole and a decade lower than the zero, we find that a dominant-pole situation almost exists and $\omega_H \simeq 10^4$ rad/s. A better estimate of ω_H can be obtained using Eq. (10.77), as follows:

$$\omega_H = 1\Bigg/\sqrt{\frac{1}{10^8}+\frac{1}{16\times 10^8}-\frac{2}{10^{10}}}$$

$$= 9800 \text{ rad/s}$$

The exact value of ω_H can be determined from the given transfer function as 9537 rad/s. Finally, we show in Fig. 10.23 a Bode plot and an exact plot for the given transfer function. Note that this is a plot of the high-frequency response of the amplifier normalized relative to its midband gain. That is, if the midband gain is, say, 100 dB, then the entire plot should be shifted upward by 100 dB.

Example 10.7 *continued*

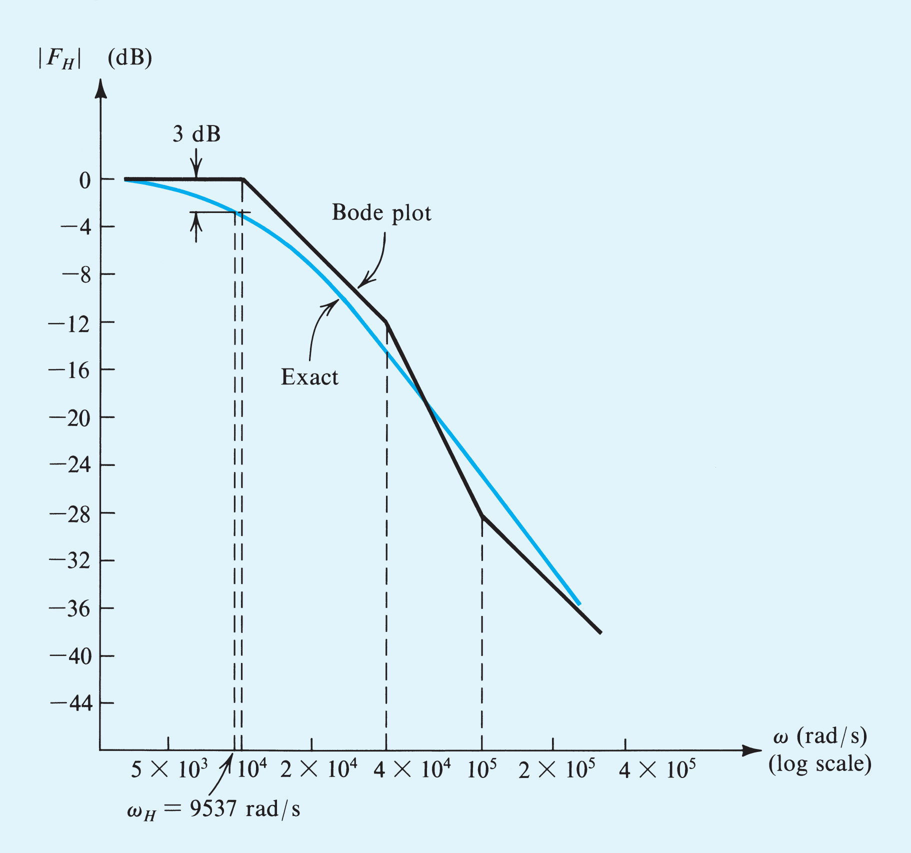


Figure 10.23 Normalized high-frequency response of the amplifier in Example 10.7.

EXERCISES

10.12 A direct-coupled amplifier has a dc gain of 1000 V/V and an upper 3-dB frequency of 100 kHz. Find the transfer function and the gain–bandwidth product in hertz.

Ans. $\dfrac{1000}{1+\dfrac{s}{2\pi\times10^5}}$; 10^8 Hz

10.13 The high-frequency response of an amplifier is characterized by two zeros at $s=\infty$ and two poles at ω_{P1} and ω_{P2}. For $\omega_{P2}=k\omega_{P1}$, find the value of k that results in the exact value of ω_H being 0.9 ω_{P1}. Repeat for $\omega_H=0.99\omega_{P1}$.

Ans. 2.78; 9.88

10.14 For the amplifier described in Exercise 10.13, find the exact and approximate values (using Eq. 10.77) of ω_H (as a function of ω_{P1}) for the cases $k=1,2$, and 4.

Ans. 0.64, 0.71; 0.84, 0.89; 0.95, 0.97

10.4.3 The Method of Open-Circuit Time Constants

If the poles and zeros of the amplifier transfer function can be determined easily, then we can determine f_H using the techniques above. In many cases, however, it is not a simple matter to determine the poles and zeros by quick hand analysis. In such cases an approximate value for f_H can be obtained using the method of open-circuit time constants, which we describe next. This method is the dual of the method of short-circuit time constants we used in Section 10.1 for the estimation of f_L.

Consider the function $F_H(s)$ (Eq. 10.71), which determines the high-frequency response of the amplifier. The numerator and denominator factors can be multiplied out and $F_H(s)$ expressed in the alternative form

$$F_H(s) = \frac{1 + a_1 s + a_2 s^2 + \cdots + a_n s^n}{1 + b_1 s + b_2 s^2 + \cdots + b_n s^n} \tag{10.78}$$

where the coefficients a and b are related to the frequencies of the zeros and poles, respectively. Specifically, the coefficient b_1 is given by

$$b_1 = \frac{1}{\omega_{P1}} + \frac{1}{\omega_{P2}} + \cdots + \frac{1}{\omega_{Pn}} \tag{10.79}$$

It can be shown [see Gray and Searle (1969)[7]] that the value of b_1 can be obtained by setting the input signal to zero and considering the various capacitances in the high-frequency equivalent circuit one at a time while reducing all other capacitors to zero (or, equivalently, replacing them with open circuits). That is, to obtain the contribution of capacitance C_i, we reduce all other capacitances to zero, reduce the input signal source to zero, and determine the resistance R_i seen by C_i. This can be done either by inspection or by replacing C_i with a voltage source V_x, finding the current I_x drawn from V_x, and calculating $R_i \equiv V_x/I_x$. The time constant τ_i associated with C_i is then found as $\tau_i = C_i R_i$. This process is then repeated for all other capacitors in the circuit. The value of b_1 is computed by summing the individual time constants, called **open-circuit time constants**,

$$b_1 = \sum_{i=1}^{n} C_i R_i \tag{10.80}$$

where we have assumed that there are n capacitors in the high-frequency equivalent circuit.

This method for determining b_1 is *exact*; the approximation comes about in using the value of b_1 to determine ω_H. Specifically, if the zeros are not dominant and if one of the poles, say P_1, is dominant, then from Eq. (10.79),

$$b_1 \simeq \frac{1}{\omega_{P1}} \tag{10.81}$$

But, also, the upper 3-dB frequency will be approximately equal to ω_{P1}, leading to

$$\omega_H \simeq \frac{1}{b_1} = \frac{1}{\sum_i C_i R_i} \tag{10.82}$$

Here it should be pointed out that in complex circuits we usually do not know whether a dominant pole exists. Nevertheless, using Eq. (10.82) to determine ω_H normally yields

[7]The bibliography is in Appendix I.

remarkably good results[8] even if a dominant pole does not exist. Finally, we note that we will sometimes refer to the sum of time constants in Eq. (10.81) by τ_H; that is, $\tau_H = b_1$, and is known as the effective high-frequency time constant.

10.4.4 Application of the Method of Open-Circuit Time Constants to the CS Amplifier

Figure 10.24 shows a generalized high-frequency equivalent circuit for the common-source amplifier. Here, V'_{sig} and R'_{sig} are the Thévenin equivalent of the signal generator together with whatever bias circuit may be present at the amplifier input [e.g., R_G in the circuit of Fig. 10.3(a)]. Resistance R'_L represents the total resistance between the output (drain) node and ground and includes R_D, r_o, and R_L (if one is present). Similarly, C_L represents the total capacitance between the drain node and ground and includes the MOSFET's drain-to-body capacitance (C_{db}), the capacitance introduced by a current-source load, the input capacitance of a succeeding amplifier stage (if one is present), and in some cases, as we will see in later chapters, a deliberately introduced capacitance. In IC MOS amplifiers, C_L can be substantial.

The equivalent circuit in Fig. 10.24 can also be used to represent the CE amplifier. Thus, we will not need to repeat the analysis; rather, we can adapt the CS results to the CE case by simply renaming the components (i.e., replacing C_{gs} by C_π and C_{gd} by C_μ).

We wish to determine the 3-dB frequency f_H of the CS amplifier in Fig. 10.24 using the method of open-circuit time constants. Toward that end, we set $V_{sig} = 0$ and consider each of the three capacitances at a time, setting the other two to zero. Figure 10.25(a) shows the resulting circuit for determining the resistance R_{gs} seen by C_{gs}, thus

$$R_{gs} = R'_{sig} \tag{10.83}$$

For C_{gd}, we obtain the circuit in Fig. 10.25(b). This circuit is somewhat complex, and we cannot determine R_{gd} by inspection. Rather, we apply a test current source I_x and determine the voltage V_x that results across I_x. We see that

$$V_{gs} = -I_x R'_{sig} \tag{10.84}$$

A loop equation gives

$$V_d = V_x + V_{gs}$$

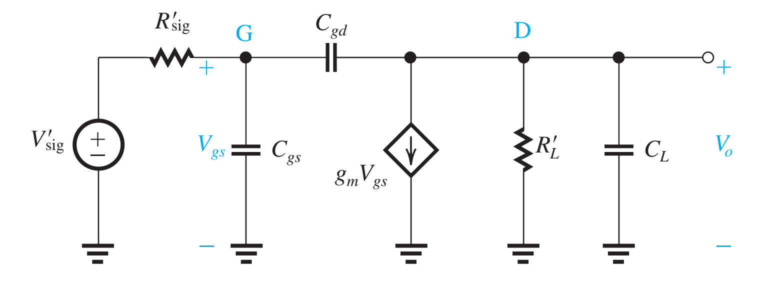


Figure 10.24 Generalized high-frequency equivalent circuit for the CS amplifier.

[8] The method of open-circuit time constants yields good results only when all the poles are real, as is the case in most of this chapter.

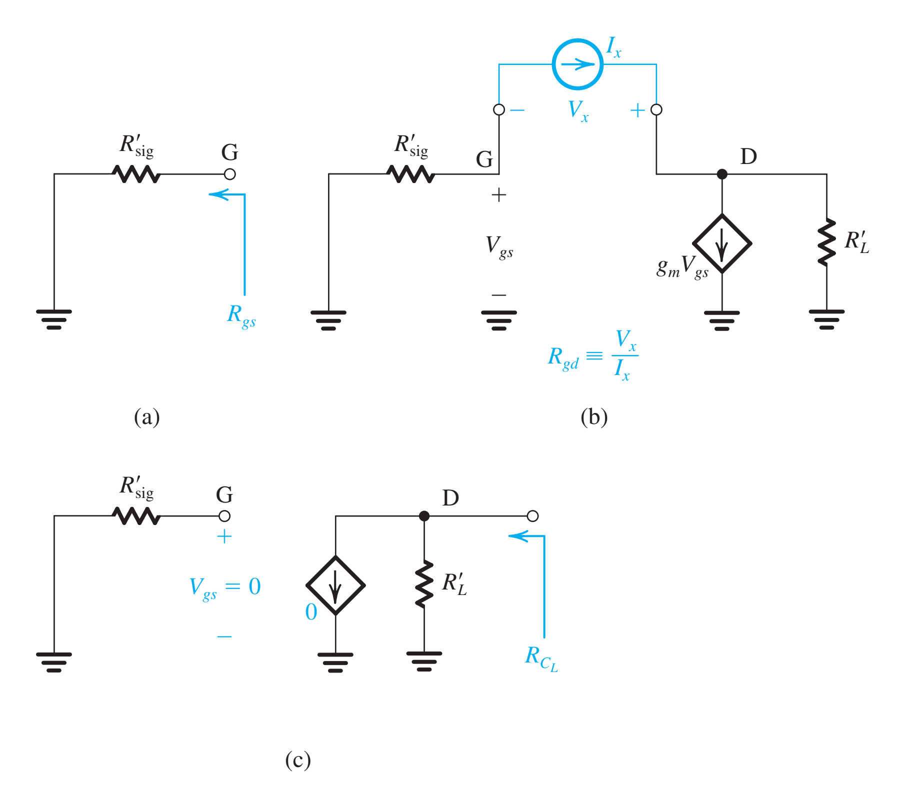


Figure 10.25 Application of the open-circuit time-constants method to the CS equivalent circuit of Fig. 10.24.

A node equation at D gives

$$I_x = g_m V_{gs} + \frac{V_d}{R'_L}$$

Thus,

$$I_x = g_m V_{gs} + \frac{V_x + V_{gs}}{R'_L}$$

Substituting for V_{gs} from Eq. (10.84) and manipulating the resulting equation gives

$$R_{gd} \equiv \frac{V_x}{I_x} = R'_{sig}(1 + g_m R'_L) + R'_L \tag{10.85}$$

Finally, for C_L we obtain the circuit shown in Fig. 10.25(c), from which

$$R_{C_L} = R'_L \tag{10.86}$$

Next, we use the resistance values in Eqs. (10.83), (10.85), and (10.86) to obtain the effective high-frequency time constant τ_H,

$$\begin{aligned} \tau_H = b_1 &= C_{gs}R_{gs} + C_{gd}R_{gd} + C_L R_{C_L} \qquad (10.87) \\ &= C_{gs}R'_{sig} + C_{gd}[R'_{sig}(1 + g_m R'_L) + R'_L] + C_L R'_L \end{aligned}$$

and the 3-dB frequency f_H is

$$f_H = \frac{1}{2\pi\tau_H}$$

An important observation is available from Eq. (10.87), which can be rewritten as

$$\tau_H = [C_{gs} + C_{gd}(1 + g_m R_L')]R_{\text{sig}}' + (C_{gd} + C_L)R_L' \tag{10.88}$$

We note that the first term is simply $C_{\text{in}}R_{\text{sig}}'$, where C_{in} is dominated by the Miller capacitance $C_{gd}(1 + g_m R_L')$. The method of open-circuit time constants, however, also provides the second term, which results from the interaction of $(C_{gd} + C_L)$ with R_L'. Thus, while the first term in Eq. (10.88) arises at the input node, the second term occurs at the output node. The second term can be dominant in cases where R_{sig}' is small, as we have seen in Section 10.3.4.

Example 10.8

An integrated-circuit CS amplifier has $g_m = 1.25$ mA/V, $C_{gs} = 20$ fF, $C_{gd} = 5$ fF, $C_L = 25$ fF, $R_{\text{sig}}' = 10\,\text{k}\Omega$, and $R_L' = 10\,\text{k}\Omega$. Determine f_H and the frequency of the transmission zero f_Z caused by C_{gd}.

Solution

To obtain f_H we first use Eqs. (10.83), (10.85), and (10.86) to determine the resistances seen by the three capacitors C_{gs}, C_{gd}, and C_L, respectively,

$$\begin{aligned}
R_{gs} &= R_{\text{sig}}' = 10\,\text{k}\Omega \\
R_{gd} &= R_{\text{sig}}'(1 + g_m R_L') + R_L' \\
&= 10(1 + 1.25 \times 10) + 10 = 145\,\text{k}\Omega \\
R_{C_L} &= R_L' = 10\,\text{k}\Omega
\end{aligned}$$

We then determine the three time constants:

$$\begin{aligned}
\tau_{gs} &= C_{gs}R_{gs} = 20 \times 10^{-15} \times 10 \times 10^3 = 200\text{ ps} \\
\tau_{gd} &= C_{gd}R_{gd} = 5 \times 10^{-15} \times 145 \times 10^3 = 725\text{ ps} \\
\tau_{C_L} &= C_L R_{C_L} = 25 \times 10^{-15} \times 10 \times 10^3 = 250\text{ ps}
\end{aligned}$$

The effective high-frequency time constant τ_H can now be obtained by summing the three time constants,

$$\begin{aligned}
\tau_H &= \tau_{gs} + \tau_{gd} + \tau_{C_L} \\
&= 200 + 725 + 250 = 1175\text{ ps}
\end{aligned}$$

and the 3-dB frequency f_H is

$$f_H = \frac{1}{2\pi\tau_H} = \frac{1}{2\pi \times 1175 \times 10^{-12}} = 135.5\text{ MHz}$$

The frequency of the transmission zero can be determined by reference to the circuit in Fig. 10.24. Since at $s = s_Z$, V_o will be zero, a node equation at the output node at $s = s_Z$ becomes

$$s_Z C_{gd}(V_{gs} - 0) = g_m V_{gs}$$

Thus

$$s_Z = \frac{g_m}{C_{gd}}$$

and

$$f_Z = \frac{g_m}{2\pi C_{gd}} = \frac{1.25 \times 10^{-3}}{2\pi \times 5 \times 10^{-15}} = 39.8 \text{ GHz}$$

which is much higher than f_H; hence it plays almost no role in the estimate of f_H.

EXERCISES

10.15 For the CS amplifier in Example 10.8, find the estimate of f_H obtained using the Miller effect method of Section 10.3.1. By what percentage does this estimate differ from that obtained in Example 10.8 using the method of open-circuit time constants? Which of the two estimates is more realistic, and why?
Ans. 181.9 MHz; greater by 34%; the value obtained by the method of open-circuit time constants is more realistic because it includes the effect of C_L.

10.16 For the CS amplifier in Example 10.8, using the value of f_H determined by the method of open-circuit time constants, find the gain–bandwidth product. Recall that $g_m = 1.25$ mA/V and $R_L' = 10\text{ k}\Omega$.
Ans. GBW = 1.69 GHz

10.17 As a way to trade gain for bandwidth, the designer of the CS amplifier in Example 10.8 connects a load resistor at the output that results in halving the value of R_L'. Find the new values of $|A_M|$, f_H, and the gain–bandwidth product.
Ans. 6.25 V/V; 223.4 MHz; 1.4 GHz

10.18 As another way to trade dc gain for bandwidth, the designer of the CS amplifier in Example 10.8 decides to operate the amplifying transistor at double the value of V_{OV} by increasing the bias current fourfold. Find the new values of g_m, R_L', $|A_M|$, f_H, and the gain–bandwidth product. Assume that R_L' is the parallel equivalent of r_o of the amplifying transistor and that of the current-source load.
Ans. 2.5 mA/V; 2.5 kΩ; 6.25 V/V; 250 MHz; 1.56 GHz

We conclude this section by emphasizing that the method of open-circuit time constants reveals to the circuit designer the relative contribution of the various capacitances to the determination of the amplifier bandwidth f_H. For instance, for the amplifier in Example 10.8, we see that while C_{gd} contributes the most (725 ps of 1175 pF, or 62%) because of the Miller effect, the contribution of each of C_{gs} (17%) and C_L (12%) is not insignificant. Such information can be useful in the amplifier-design process.

10.4.5 Application of the Method of Open-Circuit Time Constants to the CE Amplifier

The formulas developed for the CS case can be easily adapted to the case of the CE amplifier whose equivalent circuit is shown in Fig. 10.19(b). For the general case of a capacitance C_L that appears across the output terminals,

$$R_\pi = R'_{sig} \tag{10.89}$$

$$R_\mu = R'_{sig}(1 + g_m R'_L) + R'_L \tag{10.90}$$

$$R_{C_L} = R'_L \tag{10.91}$$

Thus,

$$\tau_H = C_\pi R'_{sig} + C_\mu[R'_{sig}(1 + g_m R'_L) + R'_L] + C_L R'_L \tag{10.92}$$

and

$$f_H = \frac{1}{2\pi \tau_H} \tag{10.93}$$

Note that expressions for R'_{sig} and R'_L are given in Fig. 10.19.

EXERCISE

10.19 Consider a bipolar active-loaded CE amplifier having the load current source implemented with a *pnp* transistor. Let the circuit be operating at a 1-mA bias current. The transistors are specified as follows: $\beta(npn) = 200$, $V_{An} = 130$ V, $|V_{Ap}| = 50$ V, $C_\pi = 16$ pF, $C_\mu = 0.3$ pF, $C_L = 5$ pF, and $r_x = 200\ \Omega$. The amplifier is fed with a signal source having a resistance of 36 kΩ. Determine: (a) A_M; (b) C_{in} and f_H using the Miller effect; (c) f_H using open-circuit time constants; (d) f_Z; and (e) the gain–bandwidth product.
Ans. (a) −175 V/V; (b) 450 pF, 80.6 kHz; (c) 73.5 kHz; (d) 21.2 GHz; (e) 12.9 MHz

10.5 High-Frequency Response of the Common-Gate and Cascode Amplifiers

Although common-source and common-emitter amplifiers provide substantial gain at midband frequencies, their gain falls off in the high-frequency band at a relatively low frequency. This is primarily due to the large input capacitance C_{in}, whose value is significantly increased by the Miller component. The latter is large because of the Miller multiplication effect, which the bridging capacitance C_{gd} (or C_μ) experiences. It follows that the key to obtaining wideband operation, that is, high f_H, is to use circuit configurations that do not suffer from the Miller effect. One such configuration is the common-gate circuit.

10.5.1 High-Frequency Response of the CG Amplifier

Figure 10.26(a) shows the CG amplifier with the MOSFET internal capacitances C_{gs} and C_{gd} pulled out of the model and indicated. For generality, a capacitance C_L is included at the

output node to represent the combination of the output capacitance of a current-source load and the input capacitance of a succeeding amplifier stage. Capacitance C_L also includes the MOSFET capacitance C_{db}. Note the C_L appears in effect in parallel with C_{gd}; therefore, in the following analysis we will lump the two capacitances together.

It is important to note at the outset that each of the three capacitances in the circuit of Fig. 10.26(a) has a grounded node. Therefore *none of the capacitances undergoes the Miller multiplication effect* observed in the CS stage. It follows that the CG circuit can be designed to have a much wider bandwidth than that of the CS circuit, especially when the resistance of the signal generator is large. To analyze the high-frequency response of the CG amplifier of Fig. 10.26(a), we replace the MOSFET with its T model. The resulting circuit, with C_{gd} lumped with C_L, is shown in Fig. 10.26(b).

We shall consider first the case of a discrete-circuit CG amplifier in which r_o can be neglected. Eliminating r_o results in the circuit in Fig. 10.26(c). We immediately observe that there are two poles: one at the input side with a frequency f_{P1},

$$f_{P1} = \frac{1}{2\pi C_{gs}\left(R_{\text{sig}} \,\|\, \dfrac{1}{g_m}\right)} \tag{10.94}$$

and the other at the output side with a frequency f_{P2},

$$f_{P2} = \frac{1}{2\pi(C_{\text{gd}} + C_L)R_L} \tag{10.95}$$

The relative locations of the two poles will depend on the specific situation. However, f_{P2} is usually lower than f_{P1}; thus f_{P2} can be dominant. The important point to note is that both f_{P1} and f_{P2} are usually much higher than the frequency of the dominant input pole in the CS stage. An approximate value for f_H can be obtained by applying the method of open-circuit time constants to the circuit of Fig. 10.26(c), resulting in

$$\tau_{gs} = C_{gs}\left(R_{\text{sig}} \,\|\, \frac{1}{g_m}\right) = 1/2\pi f_{P2} \tag{10.96}$$

and

$$\tau_{gd} = (C_L + C_{gd})R_L = 1/2\pi f_{P1} \tag{10.97}$$

Thus,

$$\tau_H = C_{gs}\left(R_{\text{sig}} \,\|\, \frac{1}{g_m}\right) + (C_L + C_{gd})R_L \tag{10.98}$$

and

$$f_H = \frac{1}{2\pi\tau_H} = 1\bigg/\left(\frac{1}{f_{P1}} + \frac{1}{f_{P2}}\right) \tag{10.99}$$

In IC amplifiers, r_o has to be taken into account. Applying the method of open-circuit time constants to the equivalent circuit in Fig. 10.26(b), we obtain the circuit in Fig. 10.26(d) for determining R_{gs}. From this circuit we find that

$$R_{gs} = R_{\text{sig}} \,\|\, R_{\text{in}} \tag{10.100}$$

where R_{in} is the input resistance of the CG amplifier with a load resistance R_L. An expression for R_{in} was derived in Chapter 8 and given in Eq. (8.53),

$$R_{\text{in}} = \frac{r_o + R_L}{1 + g_m r_o} \simeq \frac{r_o + R_L}{g_m r_o} \tag{10.101}$$

The resistance R_{gd} seen by $(C_L + C_{gd})$ can be obtained from the circuit in Fig. 10.26(e),

$$R_{gd} = R_L \| R_o \tag{10.102}$$

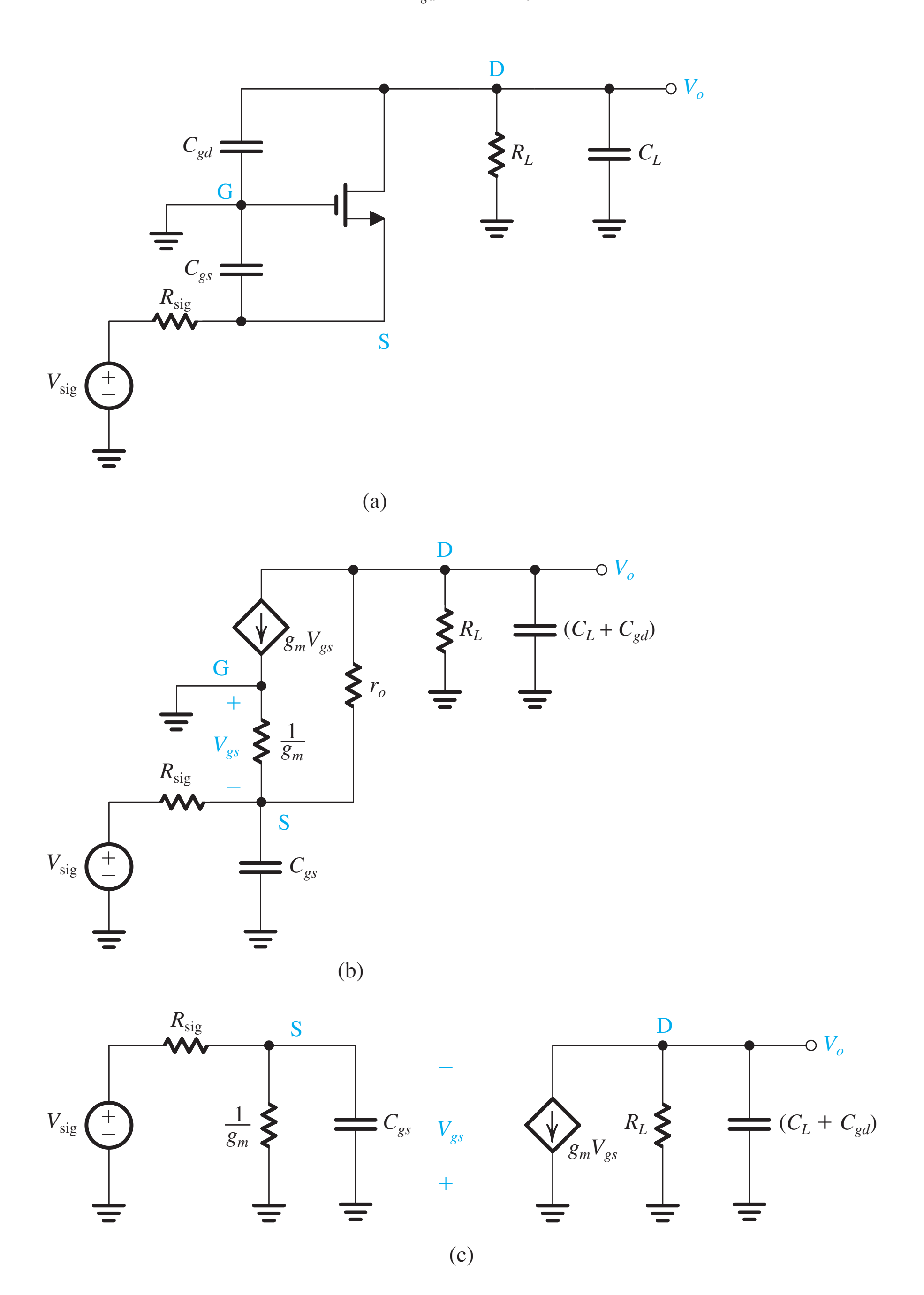


Figure 10.26 **(a)** The common-gate amplifier with the transistor internal capacitances shown. A load capacitance C_L is also included. **(b)** Equivalent circuit of the CG amplifier with the MOSFET replaced with its T model. **(c)** Equivalent circuit for the case in which r_o is neglected. **(d)** Circuit for determining the resistance R_{gs} seen by C_{gs}. **(e)** Circuit for determining the resistance R_{gd} seen by $(C_L + C_{gd})$.

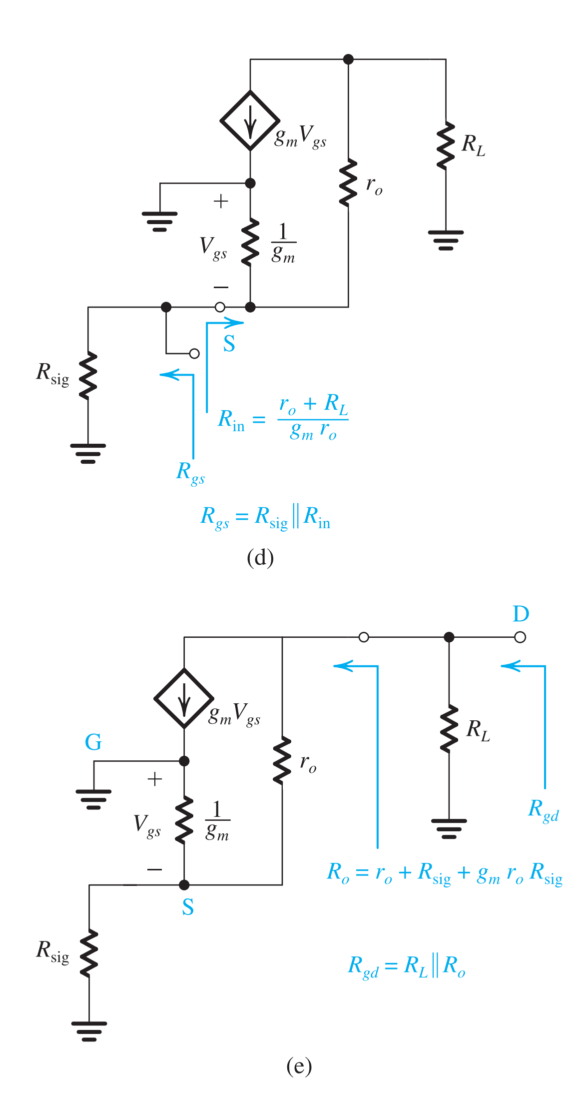


Figure 10.26 *continued*

where R_o is the output resistance of a CG amplifier with a resistance R_{sig} connected between source and ground. From Chapter 8, Eq. (8.56), we have

$$R_o = r_o + R_{sig} + g_m r_o R_{sig} \tag{10.103}$$

Finally,

$$\tau_H = \tau_{gs} + \tau_{gd} \tag{10.104}$$

and

$$f_H = \frac{1}{2\pi \tau_H} \tag{10.105}$$

Example 10.9

Consider a common-gate amplifier with $g_m = 1.25$ mA/V, $r_o = 20$ kΩ, $C_{gs} = 20$ fF, $C_{gd} = 5$ fF, $C_L = 25$ fF, $R_{sig} = 10$ kΩ, and $R_L = 20$ kΩ. Assume that C_L includes C_{db}. Determine the input resistance, the midband gain, and the upper 3-dB frequency f_H.

Solution

Figure 10.27 shows the CG amplifier circuit at midband frequencies. We note that

$$v_o = iR_L$$
$$v_{sig} = i(R_{sig} + R_{in})$$

Thus, the overall voltage gain is given by

$$G_v = \frac{v_o}{v_{sig}} = \frac{R_L}{R_{sig} + R_{in}}$$

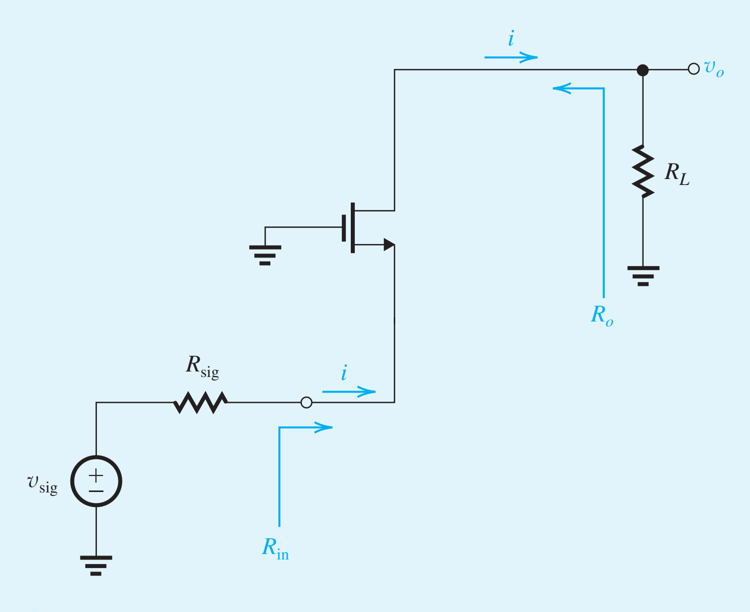


Figure 10.27 The CG amplifier circuit at midband.

The value of R_{in} is found from Eq. (10.101) as

$$R_{in} = \frac{r_o + R_L}{1 + g_m r_o}$$

$$= \frac{20 + 20}{1 + (1.25 \times 20)} = 1.54 \text{ k}\Omega$$

Thus, G_v can now be determined as

$$G_v = \frac{20}{10 + 1.54} = 1.73 \text{ V/V}$$

Observe that as expected G_v is very low. This is due to the fact that the CG amplifier draws a large input current, equal, in fact, to the load current i.

To obtain an estimate of the 3-dB frequency f_H, we first determine R_{gs} and R_{gd} using Eqs. (10.100) and (10.102),

$$R_{gs} = R_{\text{sig}} \parallel R_{\text{in}} = 10 \parallel 1.54 = 1.33 \text{ k}\Omega$$

$$R_{gd} = R_L \parallel R_o$$

where R_o is given by Eq. (10.103),

$$R_o = r_o + R_{\text{sig}} + (g_m r_o) R_{\text{sig}}$$
$$= 20 + 10 + 25 \times 10 = 280 \text{ k}\Omega$$

Thus,

$$R_{gd} = 20 \parallel 280 = 18.7 \text{ k}\Omega$$

Now we can compute the sum of the open-circuit time constants, τ_H,

$$\tau_H = C_{gs} R_{gs} + (C_{gd} + C_L) R_{gd}$$
$$\tau_H = 20 \times 10^{-15} \times 1.33 \times 10^3 + (5 + 25) \times 10^{-15} \times 18.7 \times 10^3$$
$$= 26.6 \times 10^{-12} + 561 \times 10^{-12}$$
$$= 587.6 \text{ ps}$$

and the upper 3-dB frequency f_H can be obtained as

$$f_H = \frac{1}{2\pi \tau_H} = \frac{1}{2\pi \times 587.6 \times 10^{-12}} = 270.9 \text{ MHz}$$

Observe that f_H is indeed much higher than (about twice) the corresponding value for the CS amplifier found in Example 10.8. Another important observation can be made by examining the two components of τ_H: The contribution of the input circuit is 26.6 ps, while that of the output circuit is 561 ps; thus the limitation on the high-frequency response is posed by the output circuit.

EXERCISE

10.20 In order to raise the midband gain of the CG amplifier in Example 10.9, the circuit designer decides to use a cascode current source for the load device, thus raising R_L by a factor of $g_m r_o = 25$; that is, R_L becomes 500 kΩ. Find R_{in}, the midband gain, and f_H. Comment on the results.
Ans. 20 kΩ; 16.7 V/V; 28.8 MHz. While the midband gain has been increased substantially (by a factor of 9.7), the bandwidth f_H has been substantially lowered (by a factor of about 9.4). Thus, the high-frequency advantage of the CG amplifier is completely lost!

We conclude this section by noting that a properly designed CG circuit can have a wide bandwidth. However, the input resistance will be low and the overall midband gain can be very low. It follows that the CG circuit alone will not do the job! However, combining the CG with the CS amplifier in the cascode configuration can result in a circuit having the high input resistance and gain of the CS amplifier together with the wide bandwidth of the CG amplifier, as we shall now see.

10.5.2 High-Frequency Response of the MOS Cascode Amplifier

In Section 8.5 we studied the cascode amplifier and analyzed its performance at midband frequencies. There we learned that by combining the CS and CG configurations, the cascode amplifier exhibits a very high input resistance and a voltage gain that can be as high as A_0^2, where $A_0 = g_m r_o$ is the intrinsic gain of the MOSFET. For our purposes here, we shall see that the versatility of the cascode circuit allows us to trade off some of this high midband gain in return for a wider bandwidth.

Figure 10.28 shows the cascode amplifier with all transistor internal capacitances indicated. Also included is a capacitance C_L at the output node to represent the combination of C_{db2}, the output capacitance of a current-source load, and the input capacitance of a succeeding amplifier stage (if any). Note that C_{db1} and C_{gs2} appear in parallel, and we shall combine them in the following analysis. Similarly, C_L and C_{gd2} appear in parallel and will be combined.

The easiest and, in fact, quite insightful approach to determining the 3-dB frequency f_H is to employ the open-circuit time-constants method:

1. Capacitance C_{gs1} sees a resistance R_{sig}.
2. Capacitance C_{gd1} is the gate-to-drain capacitance of the CS amplifier Q_1; thus it sees a resistance R_{gd1}, which can be obtained by adapting the formula in Eq. (10.85),

$$R_{gd1} = (1 + g_{m1}R_{d1})R_{sig} + R_{d1} \tag{10.106}$$

where R_{d1}, the total resistance at D_1, is given by the parallel equivalent of the resistance looking into the drain of Q_1 (r_{o1}) and the resistance looking into the source of Q_2 (R_{in2}),

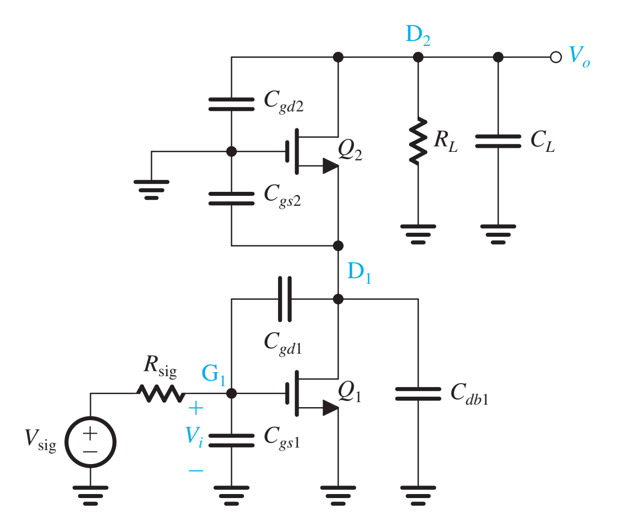


Figure 10.28 The cascode circuit with the various transistor capacitances indicated.

thus

$$R_{d1} = r_{o1} \| R_{in2} = r_{o1} \| \frac{r_{o2} + R_L}{g_{m2} r_{o2}} \tag{10.107}$$

3. Capacitance $(C_{db1} + C_{gs2})$ sees a resistance R_{d1}.
4. Capacitance $(C_L + C_{gd2})$ sees a resistance $(R_L \| R_o)$ where R_o, the output resistance of the cascode amplifier, is given by

$$R_o = r_{o2} + r_{o1} + (g_{m2} r_{o2}) r_{o1}$$

With the resistances determined, the effective time constant τ_H can be computed as

$$\tau_H = C_{gs1} R_{sig} + C_{gd1}\left[(1 + g_{m1} R_{d1}) R_{sig} + R_{d1}\right] + (C_{db1} + C_{gs2}) R_{d1} + (C_L + C_{gd2})(R_L \| R_o) \tag{10.108}$$

and the 3-dB frequency f_H as

$$f_H \simeq \frac{1}{2\pi \tau_H}$$

Design Insight and Trade-Offs To gain insight regarding what limits the high-frequency gain of the MOS cascode amplifier, we rewrite Eq. (10.108) in the form

$$\tau_H = R_{sig}\left[C_{gs1} + C_{gd1}(1 + g_{m1} R_{d1})\right] + R_{d1}(C_{gd1} + C_{db1} + C_{gs2}) + (R_L \| R_o)(C_L + C_{gd2}) \tag{10.109}$$

Here we note that the first term arises at the input node, the second term at the middle node, namely (D_1, S_2), and the third term at the output node. The first term is simply due to the interaction of the signal-source resistance R_{sig} and the input capacitance of Q_1, which, as expected, includes the Miller capacitance $C_{gd1}(1 + g_{m1} R_{d1})$.

In the case of a large R_{sig}, the first term can dominate, especially if the Miller multiplier $(1 + g_{m1} R_{d1})$ is large. This in turn happens when the load resistance R_L is large (on the order of $A_0 r_o$), causing R_{in2} and hence R_{d1} to be large and requiring the first stage, Q_1, to provide a large proportion of the gain (see Section 8.5.3). It follows that when R_{sig} is large, to extend the bandwidth we have to lower R_L to the order of r_o. This in turn lowers R_{in2} and hence R_{d1} and renders the Miller effect in Q_1 insignificant. Note, however, that the dc gain of the cascode will then be A_0. Thus, while the dc gain will be the same as (or a little higher than) that achieved in a CS amplifier, the bandwidth will be greater.

In the case when R_{sig} is small, the Miller effect in Q_1 will not be of concern. A large value of R_L (on the order of $A_0 r_o$) can then be used to realize the large dc gain possible with a cascode amplifier—that is, a dc gain on the order of A_0^2. Equation (10.109) indicates that in this case the third term will usually be dominant. To pursue this point a little further, consider the case $R_{sig} = 0$, and assume that the middle term is much smaller than the third term. It follows that

$$\tau_H \simeq (C_L + C_{gd2})(R_L \| R_o) \tag{10.110}$$

and the 3-dB frequency becomes

$$f_H = \frac{1}{2\pi (C_L + C_{gd2})(R_L \| R_o)} \tag{10.111}$$

which is of the same form as the formula for the CS amplifier with $R_{sig} = 0$ (Eq. 10.67). Here, however, $(R_L \| R_o)$ is larger that R_L' by a factor of about A_0. Thus the f_H of the cascode will be lower than that of the CS amplifier by the same factor A_0. Figure 10.29 shows a sketch of the frequency response of the cascode and of the corresponding common-source amplifier. We observe that in this case, cascoding increases the dc gain by a factor A_0 while keeping the unity-gain frequency unchanged at

$$f_t \simeq \frac{1}{2\pi} \frac{g_m}{C_L + C_{gd2}} \tag{10.112}$$

	Common Source	Cascode
Circuit	V_i, r_o, R_L, C_L, V_o; $R_L' = R_L \| r_o$	V_i, $A_0 r_o$, $A_0 R_L$, C_L, V_o
DC Gain	$-g_m R_L'$	$-A_0 g_m R_L'$
f_{3dB}	$\dfrac{1}{2\pi(C_L + C_{gd})R_L'}$	$\dfrac{1}{2\pi(C_L + C_{gd})A_0 R_L'}$
f_t	$\dfrac{g_m}{2\pi(C_L + C_{gd})}$	$\dfrac{g_m}{2\pi(C_L + C_{gd})}$

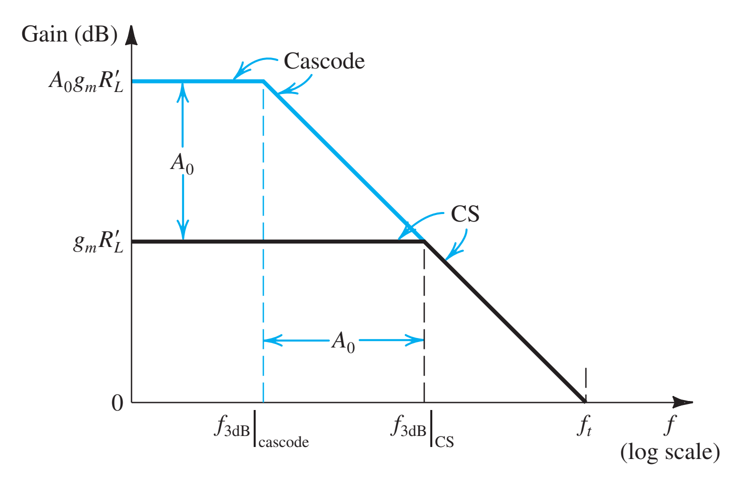


Figure 10.29 Effect of cascoding on gain and bandwidth in the case $R_{sig} = 0$. Cascoding can increase the dc gain by the factor A_0 while keeping the unity-gain frequency constant. Note that to achieve the high gain, the load resistance must be increased by the factor A_0.

Example 10.10

This example illustrates the advantages of cascoding by comparing the performance of a cascode amplifier with that of a common-source amplifier in two cases:

(a) The resistance of the signal source is significant, $R_{sig} = 10\text{ k}\Omega$.
(b) R_{sig} is negligibly small.

Assume all MOSFETs have $g_m = 1.25\text{ mA/V}$, $r_o = 20\text{ k}\Omega$, $C_{gs} = 20\text{ fF}$, $C_{gd} = 5\text{ fF}$, $C_{db} = 5\text{ fF}$, and C_L (excluding C_{db}) $= 10\text{ fF}$. For case (a), let $R_L = r_o = 20\text{ k}\Omega$ for both amplifiers. For case (b), let $R_L = r_o = 20\text{ k}\Omega$ for the CS amplifier and $R_L = R_o$ for the cascode amplifier. For all cases, determine A_v, f_H, and f_t.

Solution

(a) For the CS amplifier:

$$A_0 = g_m r_o = 1.25 \times 20 = 25\text{ V/V}$$
$$A_v = -g_m(R_L \| r_o) = -g_m(r_o \| r_o)$$
$$= -\frac{1}{2}A_0 = -12.5\text{ V/V}$$

To obtain τ_H we use Eq. (10.87) and note that $R_{sig} = R'_{sig}$ and that here C_L does not include C_{db}, thus

$$\tau_H = C_{gs}R_{sig} + C_{gd}\left[(1 + g_m R'_L)R_{sig} + R'_L\right] + (C_L + C_{db})R'_L$$

where

$$R'_L = r_o \| R_L = r_o \| r_o = 10\text{ k}\Omega$$
$$\tau_H = 20 \times 10 + 5[(1 + 12.5)10 + 10] + (10 + 5)10$$
$$= 200 + 725 + 150 = 1075\text{ ps}$$

Thus,

$$f_H = \frac{1}{2\pi \times 1075 \times 10^{-12}} = 148\text{ MHz}$$
$$f_t = |A_v| f_H = 12.5 \times 148 = 1.85\text{ GHz}$$

For the cascode amplifier:

$$R_o = 2r_o + (g_m r_o)r_o = (2 \times 20) + (25 \times 20) = 540\text{ k}\Omega$$
$$A_v = -g_m(R_o \| R_L)$$
$$= -1.25(540 \| 20) = -24.1\text{ V/V}$$
$$R_{in2} = \frac{r_o + R_L}{g_m r_o} = \frac{r_o + r_o}{g_m r_o} = \frac{2}{g_m} = \frac{2}{1.25} = 1.6\text{ k}\Omega$$
$$R_{d1} = r_o \| R_{in2} = 20 \| 1.6 = 1.48\text{ k}\Omega$$

Example 10.10 *continued*

Using Eq. (10.109),

$$\tau_H = R_{\text{sig}}\left[C_{gs1} + C_{gd1}\left(1 + g_{m1}R_{d1}\right)\right]$$
$$+ R_{d1}\left(C_{gd1} + C_{db1} + C_{gs2}\right)$$
$$+ \left(R_L \| R_o\right)\left(C_L + C_{db2} + C_{gd2}\right)$$
$$= 10[20 + 5(1 + 1.25 \times 1.48)]$$
$$+ 1.48(5 + 5 + 20)$$
$$+ (20 \| 540)(10 + 5 + 5)$$
$$= 342.5 + 44.4 + 385.7$$
$$= 772.6 \text{ ps}$$
$$f_H = \frac{1}{2\pi \times 772.6 \times 10^{-12}} = 206 \text{ MHz}$$
$$f_t = 24.1 \times 206 = 4.96 \text{ GHz}$$

Thus cascoding has increased both the dc gain and the 3-dB frequency, with the combined effect being an increase of f_t by a factor of 2.7.

(b) For the CS amplifier:

$$A_v = -12.5 \text{ V/V}$$
$$\tau_H = \left(C_{gd} + C_L + C_{db}\right)R_L'$$
$$= (5 + 10 + 5)10 = 200 \text{ ps}$$
$$f_H = \frac{1}{2\pi \times 200 \times 10^{-12}} = 796 \text{ MHz}$$
$$f_t = 12.5 \times 796 = 9.95 \text{ GHz}$$

For the cascode amplifier:

$$R_L = R_o = 540 \text{ k}\Omega$$
$$A_v = -g_m\left(R_o \| R_L\right)$$
$$= -1.25(540 \| 540) = -337.5 \text{ V/V}$$
$$R_{\text{in2}} = \frac{r_o + R_L}{g_m r_o} = \frac{20 + 540}{1.25 \times 20} = 22.4 \text{ k}\Omega$$
$$R_{d1} = r_{o1} \| R_{\text{in2}} = 20 \| 22.4 = 10.6 \text{ k}\Omega$$

$$\tau_H = R_{d1}(C_{gd1} + C_{db1} + C_{gs2}) + (R_L \| R_o)(C_L + C_{gd2} + C_{db2})$$
$$= 10.6(5 + 5 + 20) + (540 \| 540)(10 + 5 + 5)$$
$$= 318 + 5400 = 5718 \text{ ps}$$
$$f_H = \frac{1}{2\pi \times 5718 \times 10^{-12}} = 27.8 \text{ MHz}$$
$$f_t = 337.5 \times 27.8 = 9.39 \text{ GHz}$$

Thus cascoding increases the dc gain from 12.5 V/V to 337.5 V/V. This increase has been obtained at the cost of a decrease in f_H by approximately the same factor, resulting in the unity-gain frequency (which, in this case, is equal to the gain–bandwidth product) remaining nearly constant.

EXERCISE

10.21 In this exercise we wish to contrast the gain and bandwidth of a CS amplifier and a cascode amplifier. Assume that both are fed with a large source resistance R_{sig} that effectively determines the high-frequency response. Thus, neglect components of τ_H that do not include R_{sig}. Also assume that all transistors are operated at the same conditions and thus corresponding small-signal parameters are equal. Also, both amplifiers have equal $R_L = r_o$, and $g_m r_o = 40$.

(a) Find the ratio of the low-frequency gain of the cascode amplifier to that of the CS amplifier.
(b) For the case of $C_{gd} = 0.25C_{gs}$, find the ratio of f_H of the cascode to that of the CS amplifier.
(c) Use (a) and (b) to find the ratio of f_t of the cascode to that of the CS.

Ans. 2; 3.6; 7.2

Conclusion The MOS cascode amplifier is a versatile circuit that, depending on the application at hand, can be designed to provide either higher dc gain, larger bandwidth, or a combination of both, than the CS amplifier.

10.5.3 High-Frequency Response of the Bipolar Cascode Amplifier

The analysis method studied in the previous section can be directly applied to the BJT cascode amplifier. Figure 10.30 presents the circuits and the formulas for determining the high-frequency response of the bipolar cascode. Note that some of these formulas rely on the study of the bipolar cascode in Section 8.5.6.

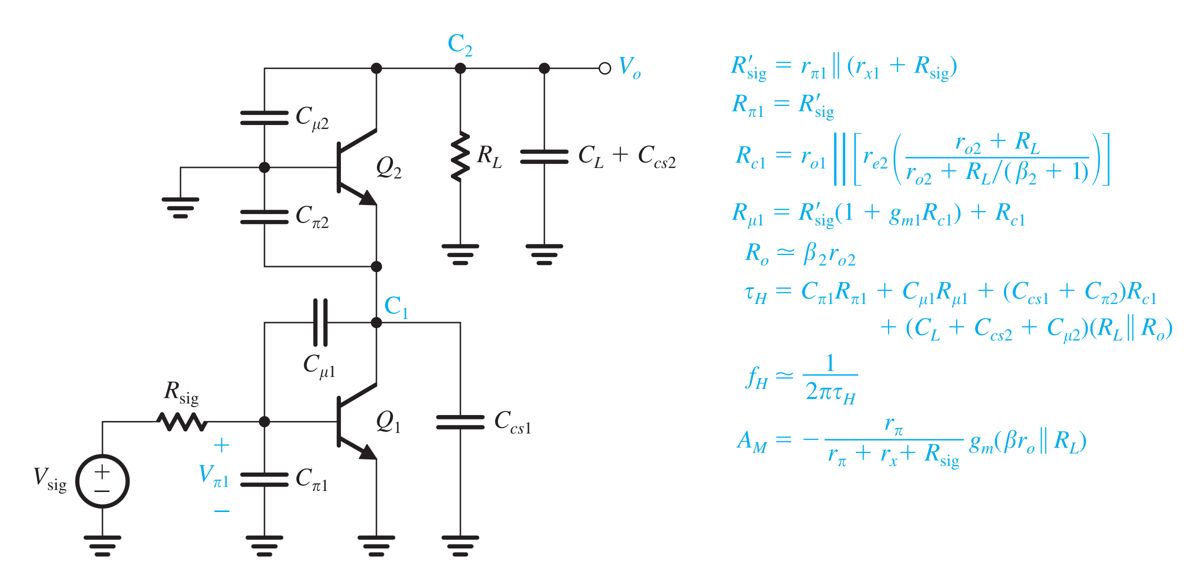


Figure 10.30 Determining the frequency response of the BJT cascode amplifier. Note that in addition to the BJT capacitances C_π and C_μ, the capacitance between the collector and the substrate C_{cs} for each transistor are included.

EXERCISE

10.22 The objective of this exercise is to evaluate the effect of cascoding on the performance of the CE amplifier of Exercise 10.19. The specifications are as follows: $I = 1$ mA, $\beta = 200$, $r_o = 130$ kΩ, $C_\pi = 16$ pF, $C_\mu = 0.3$ pF, $r_x = 200$ Ω, $C_{cs1} = C_{cs2} = 0$, $C_L = 5$ pF, $R_{sig} = 36$ kΩ, $R_L = 50$ kΩ. Find R_{in}, A_0, R_{o1}, R_{in2}, R_o, A_M, f_H, and f_t. Compare A_M, f_H, and f_t with the corresponding values obtained in Exercise 10.19 for the CE amplifier. What should C_L be reduced to in order to have $f_H = 1$ MHz?
Ans. 5.2 kΩ; 5200 V/V; 130 kΩ; 35 Ω; 26 MΩ; −242 V/V; 470 kHz; 113.8 MHz. $|A_M|$ has increased from 175 V/V to 248 V/V; f_H has increased from 73.5 kHz to 470 kHz; f_t has increased from 12.9 MHz to 113.8 MHz. C_L must be reduced to 1.6 pF.

A 10.6 High-Frequency Response of the Source and Emitter Followers

In this section, we study the high-frequency response of two important circuit building blocks: the source follower and the emitter follower. Both have a midband voltage gain that is less than but close to unity. Their advantage lies in their high input resistance and low output resistance. Thus, they find application as the output stage of a multistage amplifier and as a voltage buffer. As will be seen shortly, these voltage followers have another important advantage, namely, a wide bandwidth.

Analysis of the high-frequency response of the source and emitter followers is somewhat involved. This is because the follower has two high-frequency poles that can be close to each

other on the negative real axis of the s plane. Furthermore, in many cases, the poles can become complex. As a result, the method of open-circuit time constants cannot be used to determine f_H of the followers except in special circumstances. Our approach, therefore, will be to analyze the follower circuit to determine its gain V_o/V_{sig} as a function of frequency and then use it to determine f_H. Although the analysis is somewhat lengthy, the results can be applied easily. In the following we shall do the analysis of the source follower in detail. Then, because of similarity, the results for the emitter follower will be given without proof.

10.6.1 The Source-Follower Case

Figure 10.31(a) shows a source follower without the biasing arrangement. The follower is driven by a signal source (V_{sig}, R_{sig}) and is loaded with a resistance R_L and, for generality, a capacitance C_L. Replacing the MOSFET with its hybrid-π equivalent-circuit model results in the equivalent circuit shown in Fig. 10.31(b). Here, we have included the body-effect generator $g_{mb}V_{bs}$ because it plays an important role in determining the source-follower gain. Also, we are assuming that whatever capacitances exist between the MOSFET source and ground, such as C_{sb} of Fig. 10.12(a), have been lumped into C_L.

Noting that the drain terminal is grounded, we see that C_{gd} in fact appears across the input terminals of the source follower. Also, r_o is in parallel with R_L and can be combined with it. Finally, we observe that since the body terminal B is connected to ground, the voltage V_{bs} appears across the controlled source $g_{mb}V_{bs}$. Thus we can utilize the source-absorption theorem (see Appendix D) to replace the controlled source with a resistance $1/g_{mb}$. Since the latter appears between source and ground, it is in parallel with R_L and can be combined with it.

Utilizing the above observations, we obtain the simplified equivalent circuit shown in Fig. 10.31(c) where

$$R_L' = R_L \parallel r_o \parallel \frac{1}{g_{mb}} \tag{10.113}$$

Obtaining the Transfer Function $V_o(s)/V_{\text{sig}}(s)$ Analysis of the equivalent circuit in Fig. 10.31(c) to determine the transfer function $V_o(s)/V_{\text{sig}}(s)$ proceeds as follows.

1. A loop equation at the input yields

$$V_{\text{sig}} = I_i R_{\text{sig}} + V_g$$

where V_g can be expressed as

$$V_g = V_{gs} + V_o$$

Thus,

$$V_{\text{sig}} = I_i R_{\text{sig}} + V_{gs} + V_o \tag{10.114}$$

2. A node equation at G provides

$$\begin{aligned} I_i &= sC_{gd}V_g + sC_{gs}V_{gs} \\ &= sC_{gd}(V_{gs} + V_o) + sC_{gs}V_{gs} \end{aligned}$$

which can be substituted into Eq. (10.114) to obtain

$$V_{\text{sig}} = [1 + s(C_{gs} + C_{gd})R_{\text{sig}}]V_{gs} + [1 + sC_{gd}R_{\text{sig}}]V_o \tag{10.115}$$

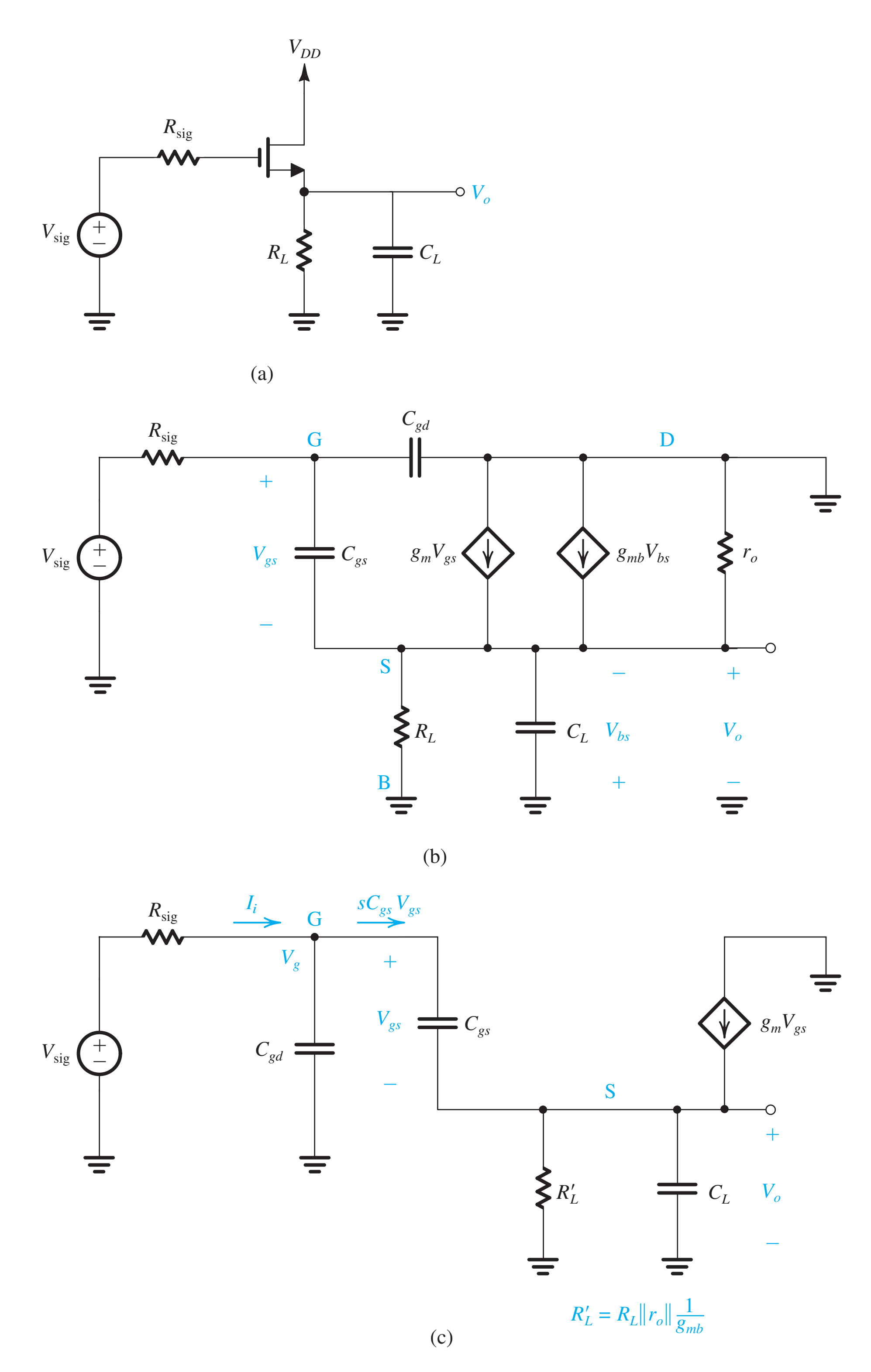


Figure 10.31 (**a**) A directly coupled source follower without the bias detail; (**b**) high-frequency equivalent circuit of the source follower; (**c**) a simplified version of the equivalent circuit.

3. A node equation at S gives

$$(g_m + sC_{gs})V_{gs} = \left(\frac{1}{R_L'} + sC_L\right)V_o$$

which can be used to express V_{gs} in terms of V_o as

$$V_{gs} = \frac{1}{g_m R_L'} \frac{1 + sC_L R_L'}{1 + s(C_{gs}/g_m)} V_o \qquad (10.116)$$

Substituting this expression of V_{gs} into Eq. (10.115) results in an equation containing only V_o and V_{sig}; this equation can be manipulated to obtain the source-follower transfer function in the form

$$\frac{V_o}{V_{sig}}(s) = A_M \frac{1 + \left(\frac{s}{\omega_Z}\right)}{1 + b_1 s^2 + b_2 s^2} \qquad (10.117)$$

where

$$A_M = \frac{R_L'}{R_L' + \frac{1}{g_m}} = \frac{g_m R_L'}{g_m R_L' + 1} \qquad (10.118)$$

$$\omega_Z = g_m / C_{gs} \qquad (10.119)$$

$$b_1 = \left(C_{gd} + \frac{C_{gs}}{g_m R_L' + 1}\right)R_{sig} + \left(\frac{C_{gs} + C_L}{g_m R_L' + 1}\right)R_L' \qquad (10.120)$$

$$b_2 = \frac{(C_{gs} + C_{gd})C_L + C_{gs}C_{gd}}{g_m R_L' + 1} R_{sig} R_L' \qquad (10.121)$$

Analysis of the Source-Follower Transfer Function We now make a number of observations on the transfer function in Eq. (10.117), which describes the gain of the source follower at high frequencies:

1. Since the source follower in Fig. 10.31(a) is directly coupled, the gain at dc is equal to A_M. This correlates with Eq. (10.117) as

$$A_M = V_o/V_{sig}\,|_{s=0}$$

2. Although the equivalent circuit of Fig. 10.31(c) has three capacitors, the transfer function is of second order. This is because the three capacitors form a continuous loop.
3. The two transmission zeros can be found from Eq. (10.117) as the values of s for which $V_o/V_{sig} = 0$. From Eq. (10.117), we see that V_o/V_{sig} approaches 0 as s approaches ∞. Thus one transmission zero is at $s = \infty$. Physically, this zero is a result of C_{gd}, which appears across the input terminals, becoming a short circuit at infinite frequency and thus making $V_o = 0$. From the numerator of Eq. (10.117) we see that the other transmission zero is at $s = -\omega_Z$ where ω_Z is given by Eq. (10.119). We note that ω_Z is slightly higher than the unity-gain frequency ω_T of the MOSFET [Eq. (10.29)],

$$\omega_T = \frac{g_m}{C_{gs} + C_{gd}} \qquad (10.122)$$

Thus the finite transmission zero is at such a high frequency that its effect on the frequency response of the follower is negligibly small.

4. The two poles of the source follower can be found as the roots of the denominator polynomial $(1+b_1s+b_2s^2)$. If the poles are real, their frequencies, say ω_{P1} and ω_{P2}, can be found from

$$1+b_1s+b_2s^2 = \left(1+\frac{s}{\omega_{P1}}\right)\left(1+\frac{s}{\omega_{P2}}\right) \tag{10.123}$$

Now if $\omega_{P2} \gg \omega_{P1}$ (at least four times larger), a dominant pole exists with frequency ω_{P1} and the 3-dB frequency f_H is given by

$$f_H \simeq f_{P1} \simeq \frac{1}{2\pi b_1} \tag{10.124}$$

Here we remind the reader that b_1 is also τ_H, the effective high-frequency time constant evaluated in the method of open-circuit time constants.

5. If the poles are real but none is dominant, the 3-dB frequency can be determined analytically from the transfer function as the frequency at which $|V_o/V_{sig}| = A_M/\sqrt{2}$. An approximate value can be obtained using the formula in Eq. (10.77),

$$f_H \simeq 1\bigg/\sqrt{\frac{1}{f_{P1}^2}+\frac{1}{f_{P2}^2}-\frac{2}{f_Z^2}} \tag{10.125}$$

6. If the poles are complex, they are best described in terms of their frequency ω_0 and Q-factor, where

$$1+b_1s+b_2s^2 = 1+\frac{1}{Q}\frac{s}{\omega_0}+\frac{s^2}{\omega_0^2} \tag{10.126}$$

and for complex poles, $Q > 0.5$. Figure 10.32(a) provides a geometrical interpretation of ω_0 and Q. From the study of second-order network responses in Chapter 17, it will be seen that the response of the source follower shows no peaking for $Q \leq 0.707$. The boundary case corresponding to $Q = 0.707$ (poles at 45° angles) results in what is known as a **maximally flat response** for which $f_{3dB} = f_0$. Figure 10.32(b) shows a number of possible responses obtained for various values of Q. In terms of the component values of the source follower,

$$\omega_0 = \frac{1}{\sqrt{b_2}} = \sqrt{\frac{g_mR_L'+1}{R_{sig}R_L'[(C_{gs}+C_{gd})C_L+C_{gs}C_{gd}]}} \tag{10.127}$$

$$Q = \frac{\sqrt{b_2}}{b_1} = \frac{\sqrt{g_mR_L'+1}\sqrt{[(C_{gs}+C_{gd})C_L+C_{gs}C_{gd}]R_{sig}R_L'}}{[C_{gs}+C_{gd}(g_mR_L'+1)]R_{sig}+(C_{gs}+C_L)R_L'} \tag{10.128}$$

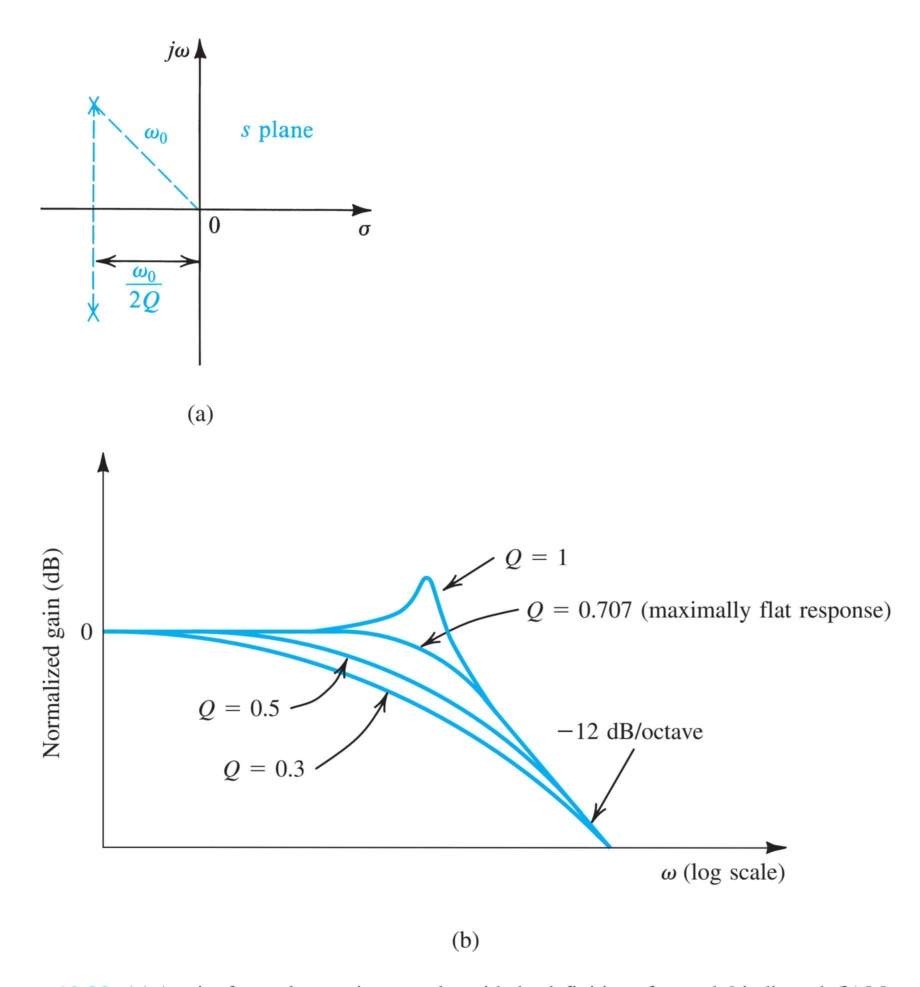


Figure 10.32 (**a**) A pair of complex-conjugate poles with the definition of ω_o and Q indicated. (**b**) Magnitude response of a source (or emitter) follower for different values of the parameter Q. Note that the response is normalized relative to A_M.

Example 10.11

A source follower operated at $g_m = 2$ mA/V and $r_o = 20\,\text{k}\Omega$ is fed with a signal source for which $R_{\text{sig}} = 10\,\text{k}\Omega$ and is loaded in a resistance $R_L = 20\,\text{k}\Omega$. The MOSFET has $C_{gs} = 20$ fF, $C_{gd} = 5$ fF, and $g_{mb} = \chi g_m$ where $\chi = 0.2$, and the total capacitance at the output $C_L = 15$ fF. Determine $A_M, f_T, f_Z, Q, f_{P1}, f_{P2}$, and $f_{3\text{dB}}$.

Solution

$$R_L' = R_L \,\|\, r_o \,\|\, \frac{1}{g_{mb}}$$

$$= 20 \,\|\, 20 \,\|\, \frac{1}{0.2 \times 2} = 20 \,\|\, 20 \,\|\, 2.5 = 2\,\text{k}\Omega$$

Example 10.11 *continued*

$$A_M = \frac{R'_L}{R'_L + \dfrac{1}{g_m}} = \frac{2}{2 + \dfrac{1}{2}} = 0.8 \text{ V/V}$$

$$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$$

$$= \frac{2 \times 10^{-3}}{2\pi(20+5) \times 10^{-15}}$$

$$= 12.7 \text{ GHz}$$

$$f_Z = \frac{g_m}{2\pi C_{gs}} = \frac{2 \times 10^{-3}}{2\pi \times 20 \times 10^{-15}} = 15.9 \text{ GHz}$$

To evaluate Q we substitute the given component values into Eq. (10.128),

$$Q = 0.42$$

Thus the poles are real. Their frequencies can be obtained by finding the roots of the polynomial $(1 + b_1 s + b_2 s^2)$, where

$$b_1 = \tau_H = 104 \text{ ps}$$

and

$$b_2 = 1.9 \times 10^{-21}$$

Thus,

$$f_{P1} = 1.98 \text{ GHz}$$

$$f_{P2} = 6.73 \text{ GHz}$$

Since $f_{P2}/f_{P1} = 3.4 < 4$, no dominant pole exists. An approximate value for f_H can be obtained as

$$f_H = 1 \Bigg/ \sqrt{\frac{1}{f_{P1}^2} + \frac{1}{f_{P2}^2} - \frac{2}{f_Z^2}} = 1.93 \text{ GHz}$$

The exact value of f_H can be found from the transfer function as 1.86 GHz, which is not much different from the approximate value.

EXERCISES

10.23 Recalling that $\tau_H = b_1$, use the expression for b_1 in Eq. (10.120) to find expressions for the three resistances R_{gs}, R_{gd}, and R_{C_L} for the source follower.

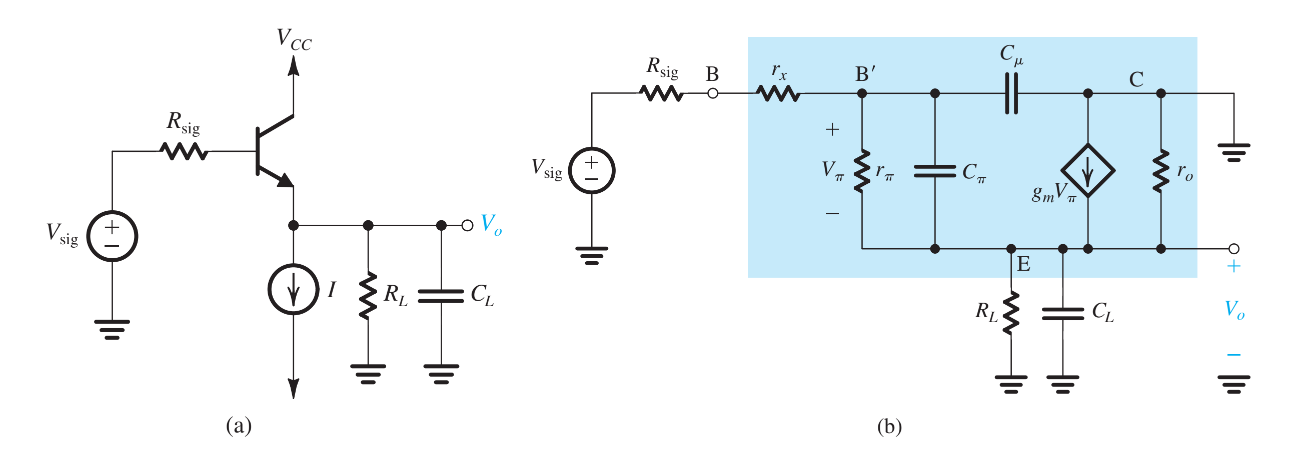
Ans. $R_{gs} = \dfrac{R_{\text{sig}} + R'_L}{g_m R'_L + 1}$; $R_{gd} = R_{\text{sig}}$; $R_{C_L} = \dfrac{R'_L}{g_m R'_L + 1}$

10.24 In Example 10.11, even though we found that a dominant pole does not exist, use the method of open-circuit time constants to obtain an estimate for f_H. (*Hint*: Recall that $\tau_H = b_1$.)

Ans. $f_H = 1.53$ GHz; about 18% lower than the exact value of 1.86 GHz; still not a bad estimate!

10.6.2 The Emitter-Follower Case

Figure 10.33 provides the results for the case of the emitter follower. The analysis here is a little more complicated because of the finite β of the BJT.



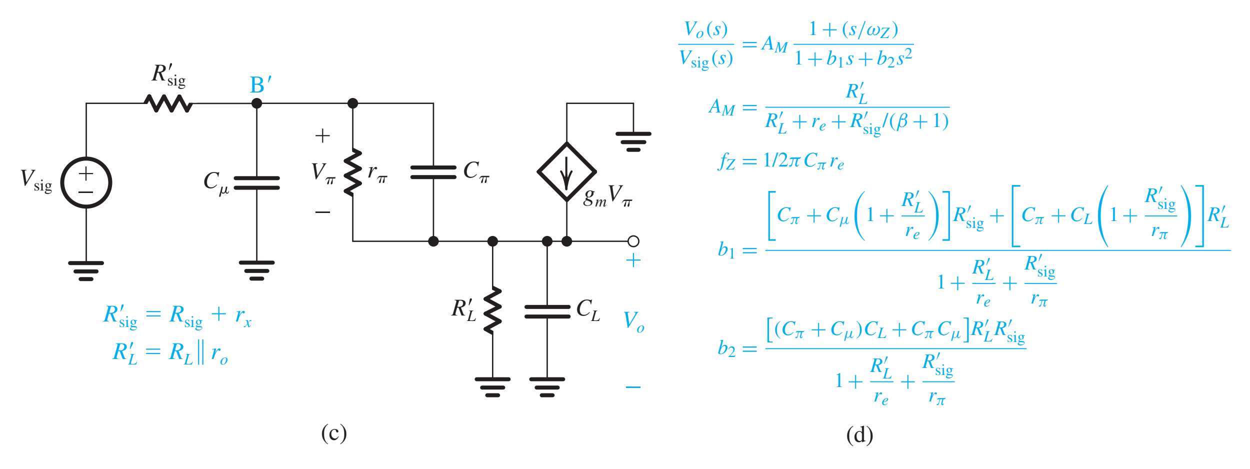


Figure 10.33 **(a)** Emitter follower. **(b)** High-frequency equivalent circuit. **(c)** Simplified equivalent circuit. **(d)** Transfer function.

EXERCISE

10.25 For an emitter follower biased at $I_C = 1$ mA and having $R_{sig} = R_L = 1$ kΩ, $r_o = 100$ kΩ, $\beta = 100$, $C_\mu = 2$ pF, $C_L = 0$, $r_x = 0$, and $f_T = 400$ MHz, find the low-frequency gain A_M, f_Z, f_{P1}, f_{P2}, and an estimate for f_H.
Ans. 0.97 V/V; 458 MHz; 67.2 MHz; 562 MHz; 67.2 MHz



10.7 High-Frequency Response of Differential Amplifiers

In this section we study the high-frequency response of the differential amplifier. We will consider the variation with frequency of both the differential gain and the common-mode gain and hence of the CMRR. We will rely heavily on the study of frequency response of single-ended amplifiers presented in the sections above. Also, we will consider MOS circuits only; the bipolar case is a straightforward extension, as we saw above on a number of occasions.

10.7.1 Analysis of the Resistively Loaded MOS Amplifier

We begin with the basic, resistively loaded MOS differential pair shown in Fig. 10.34(a). Note that we have explicitly shown transistor Q_S that supplies the bias current I. Although we are showing a dc bias voltage V_{BIAS} at its gate, usually Q_S is part of a current mirror. This detail, however, is of no consequence to our present needs. Most importantly, we are interested in the total impedance between node S and ground, Z_{SS}, because this impedance plays a significant role in determining the common-mode gain and the CMRR of the differential amplifier. Resistance R_{SS} is simply the output resistance of current source Q_S. Capacitance C_{SS} is the total capacitance between node S and ground and includes C_{db} and C_{gd} of Q_S, as well as C_{sb1} and C_{sb2}. This capacitance can be significant, especially if wide transistors are used for Q_S, Q_1, and Q_2.

The differential half-circuit shown in Fig. 10.34(b) can be used to determine the frequency dependence of the differential gain V_o/V_{id}. Indeed the gain function $A_d(s)$ of the differential

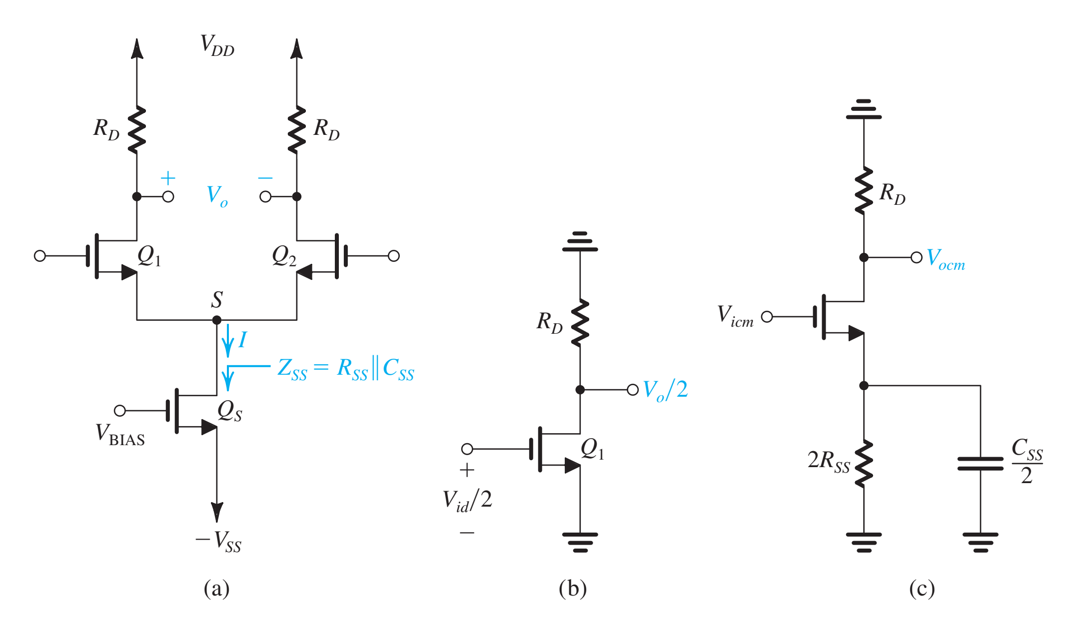


Figure 10.34 **(a)** A resistively loaded MOS differential pair; the transistor supplying the bias current is explicitly shown. It is assumed that the total impedance between node S and ground, Z_{SS}, consists of a resistance R_{SS} in parallel with a capacitance C_{SS}. **(b)** Differential half-circuit. **(c)** Common-mode half-circuit.

amplifier will be identical to the transfer function of this common-source amplifier.[9] We studied the frequency response of the common-source amplifier at great length in Sections 10.3 and 10.4 and will not repeat this material here.

EXERCISE

10.26 A MOSFET differential amplifier such as that in Fig. 10.34(a) is biased with a current $I = 0.8$ mA. The transistors Q_1 and Q_2 have $W/L = 100$, $k_n' = 0.2\ \text{mA/V}^2$, $V_A = 20$ V, $C_{gs} = 50$ fF, $C_{gd} = 10$ fF, and $C_{db} = 10$ fF. The drain resistors are 5 kΩ each. Also, there is a 100-fF capacitive load between each drain and ground.

(a) Find V_{OV} and g_m for each transistor.
(b) Find the differential gain A_d.
(c) If the input signal source has a small resistance R_{sig} and thus the frequency response is determined primarily by the output pole, estimate the 3-dB frequency f_H. [*Hint:* Refer to Section 10.3.4 and specifically to Eq. (10.67).]
(d) If, in a different situation, the amplifier is fed symmetrically with a signal source of 20 kΩ resistance (i.e., 10 kΩ in series with each gate terminal), use the open-circuit time-constants method to estimate f_H. [*Hint:* Refer to Section 10.4.4 and specifically to Eq. (10.87).]

Ans. (a) 0.2 V, 4 mA/V; (b) 18.2 V/V; (c) 292 MHz; (d) 53.7 MHz

The common-mode half-circuit is shown in Fig. 10.34(c). Although this circuit has other capacitances, namely, C_{gs}, C_{gd}, and C_{db} of the transistor in addition to other stray capacitances, we have chosen to show only $C_{SS}/2$. This is because $(C_{SS}/2)$ together with $(2R_{SS})$ form a real-axis zero in the common-mode gain function at a frequency much lower than those of the other poles and zeros of the circuit. This zero then dominates the frequency dependence of A_{cm} and CMRR.

If the output of the differential amplifier is taken single-endedly, then the common-mode gain of interest is V_{ocm}/V_{icm}. More typically, the output is taken differentially. Nevertheless, as we have seen in Section 9.3, V_{ocm}/V_{icm} still plays a major role in determining the common-mode gain. To be specific, consider what happens when the output is taken differentially and there is a mismatch ΔR_D between the two drain resistances. The resulting common-mode gain was found in Section 9.3 to be (Eq. 9.80)

$$A_{cm} = -\left(\frac{R_D}{2R_{SS}}\right)\frac{\Delta R_D}{R_D} \tag{10.129}$$

which is simply the product of V_{ocm}/V_{icm} and the per-unit mismatch $(\Delta R_D/R_D)$. Similar expressions can be found for the effects of other circuit mismatches. The important point to note is that the factor $R_D/2R_{SS}$ is always present in these expressions. Thus, the frequency dependence of A_{cm} can be obtained by simply replacing R_{SS} by Z_{SS} in this factor. Doing so for

[9] Here we are not showing the resistance of the signal source R_{sig}, which, of course, must be included in the frequency-response analysis, as we have done in the case of the CS amplifier. See Exercise 10.26.

the expression in Eq. (10.129) gives

$$A_{cm}(s) = -\frac{R_D}{2Z_{SS}}\left(\frac{\Delta R_D}{R_D}\right)$$

$$= -\frac{1}{2}R_D\left(\frac{\Delta R_D}{R_D}\right)Y_{SS}$$

$$= -\frac{1}{2}R_D\left(\frac{\Delta R_D}{R_D}\right)\left(\frac{1}{R_{SS}} + sC_{SS}\right)$$

$$= -\frac{R_D}{2R_{SS}}\left(\frac{\Delta R_D}{R_D}\right)(1 + sC_{SS}R_{SS}) \qquad (10.130)$$

from which we see that A_{cm} acquires a zero on the negative real axis of the s plane with frequency ω_Z,

$$\omega_Z = \frac{1}{C_{SS}R_{SS}} \qquad (10.131)$$

or in hertz,

$$f_Z = \frac{\omega_Z}{2\pi} = \frac{1}{2\pi C_{SS}R_{SS}} \qquad (10.132)$$

As mentioned above, usually f_Z is much lower than the frequencies of the other poles and zeros. As a result, the common-mode gain increases at the rate of +6 dB/octave (20 dB/decade) starting at a relatively low frequency, as indicated in Fig. 10.35(a). Of course, A_{cm} drops off at high frequencies because of the other poles of the common-mode half-circuit. It is, however, f_Z that is significant, for it is the frequency at which the CMRR of the differential amplifier begins to decrease, as indicated in Fig. 10.35(c). Note that if both A_d and A_{cm} are expressed and plotted in dB, then CMRR in dB is simply the difference between A_d and A_{cm}.

Although in the foregoing we considered only the common-mode gain resulting from an R_D mismatch, the results apply to the common-mode gain resulting from any other mismatch. For instance, it applies equally well to the case of a g_m mismatch, modifying Eq. (9.87) by replacing R_{SS} by Z_{SS}, and so on.

Before leaving this section, it is interesting to point out an important trade-off found in the design of the current-source transistor Q_S: In order to operate this current source with a small V_{DS} (to conserve the already low V_{DD}), we desire to operate the transistor at a low overdrive voltage V_{OV}. For a given value of the current I, however, this means using a large W/L ratio (i.e., a wide transistor). This in turn increases C_{SS} and hence lowers f_Z with the result that the CMRR deteriorates (i.e., decreases) at a relatively low frequency. Thus there is a trade-off between the need to reduce the dc voltage across Q_S and the need to keep the CMRR reasonably high at higher frequencies.

To appreciate the need for high CMRR at higher frequencies, consider the situation illustrated in Fig. 10.36: We show two stages of a differential amplifier whose power-supply voltage V_{DD} is corrupted with high-frequency noise. Since the quiescent voltage at each of the drains of Q_1 and Q_2 is $[V_{DD} - (I/2)R_D]$ we see that v_{D1} and v_{D2} will have the same high-frequency noise as V_{DD}. This high-frequency noise then constitutes a common-mode input signal to the second differential stage, formed by Q_3 and Q_4. If the second differential stage is perfectly matched, its differential output voltage V_o should be free of high-frequency

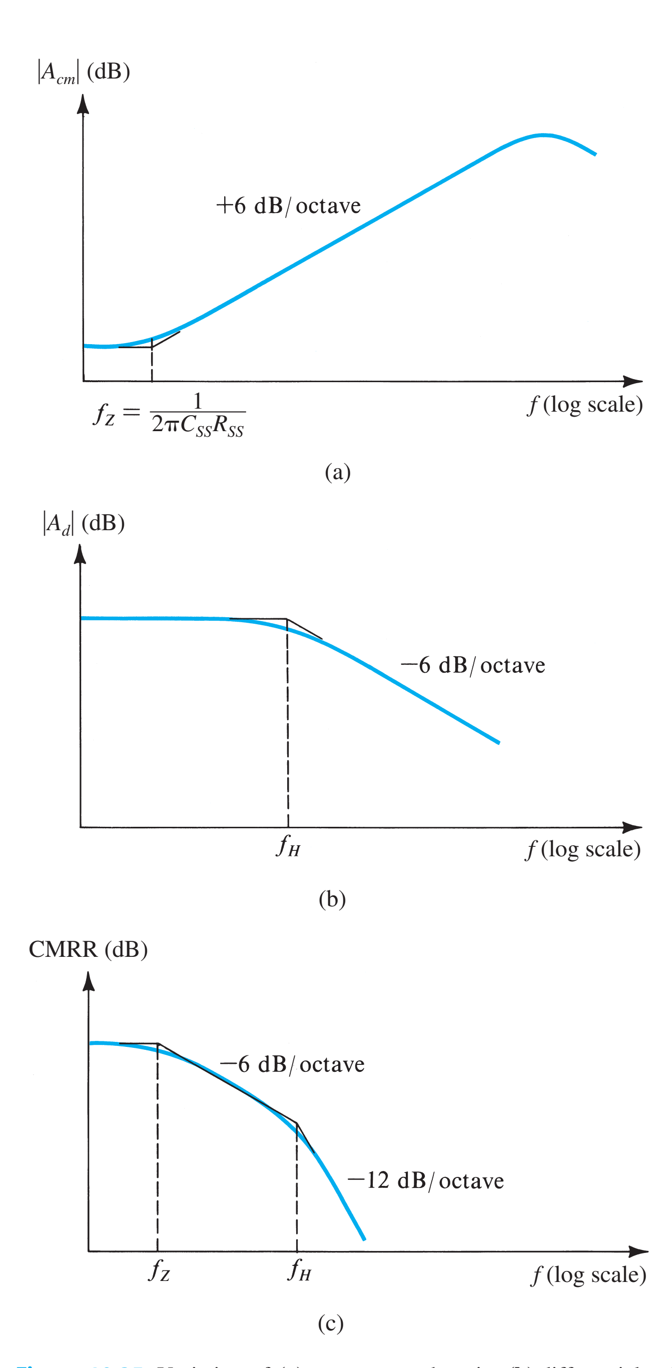


Figure 10.35 Variation of **(a)** common-mode gain, **(b)** differential gain, and **(c)** common-mode rejection ratio with frequency.

noise. However, in practice there is no such thing as perfect matching, and the second stage will have a finite common-mode gain. Furthermore, because of the zero formed by R_{SS} and C_{SS} of the second stage, the common-mode gain will increase with frequency, causing some of the noise to make its way to V_o. With careful design, this undesirable component of V_o can be kept small.

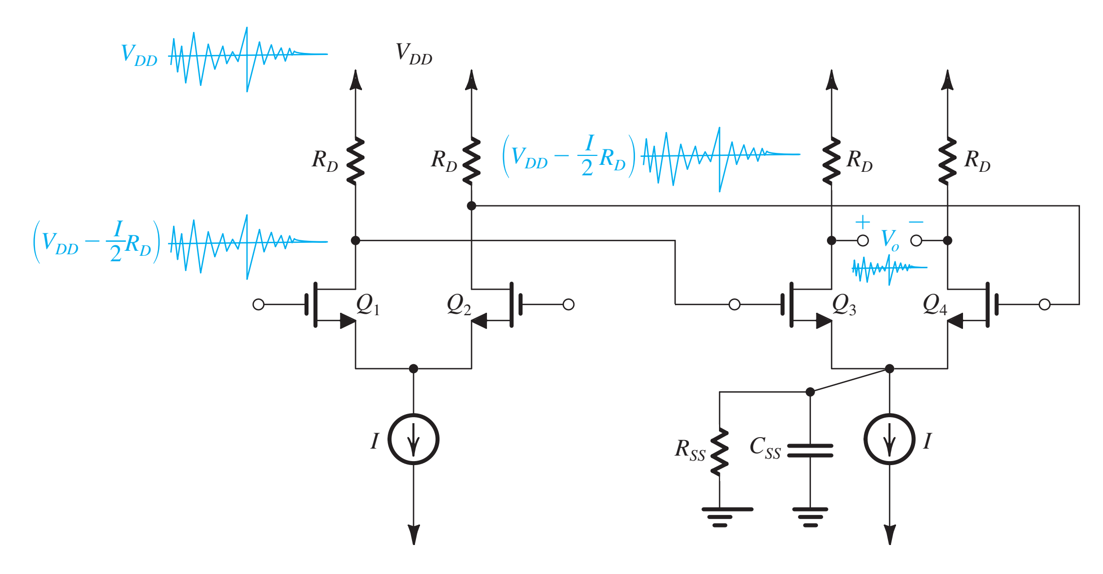


Figure 10.36 The second stage in a differential amplifier, which is relied on to suppress high-frequency noise injected by the power supply of the first stage, and therefore must maintain a high CMRR at higher frequencies.

EXERCISE

10.27 The differential amplifier specified in Exercise 10.26 has $R_{SS} = 75\text{ k}\Omega$ and $C_{SS} = 0.4\text{ pF}$. Find the 3-dB frequency of the CMRR.
Ans. 5.3 MHz

RFID—IDENTIFICATION AT A DISTANCE:

Radio-frequency identification (RFID) tags are increasingly important elements in our daily lives. These tiny devices embody a wide range of electronic techniques: linear, digital, radio, power, and signaling, all at very low power levels. They are activated (and powered) by an external interrogation signal that can be located a few feet away. Once activated, they respond with the requested data, most often simply an identifying number. Applications are incredibly diverse, ranging from automobile key fobs to tags that permit the recovery of lost pets.

10.7.2 Analysis of the Current-Mirror-Loaded MOS Amplifier

We next consider the frequency response of the current-mirror-loaded MOS differential-pair circuit studied in Section 9.5. The circuit is shown in Fig. 10.37(a) with two capacitances indicated: C_m, which is the total capacitance at the input node of the current mirror, and C_L, which is the total capacitance at the output node. Capacitance C_m is mainly formed by C_{gs3}

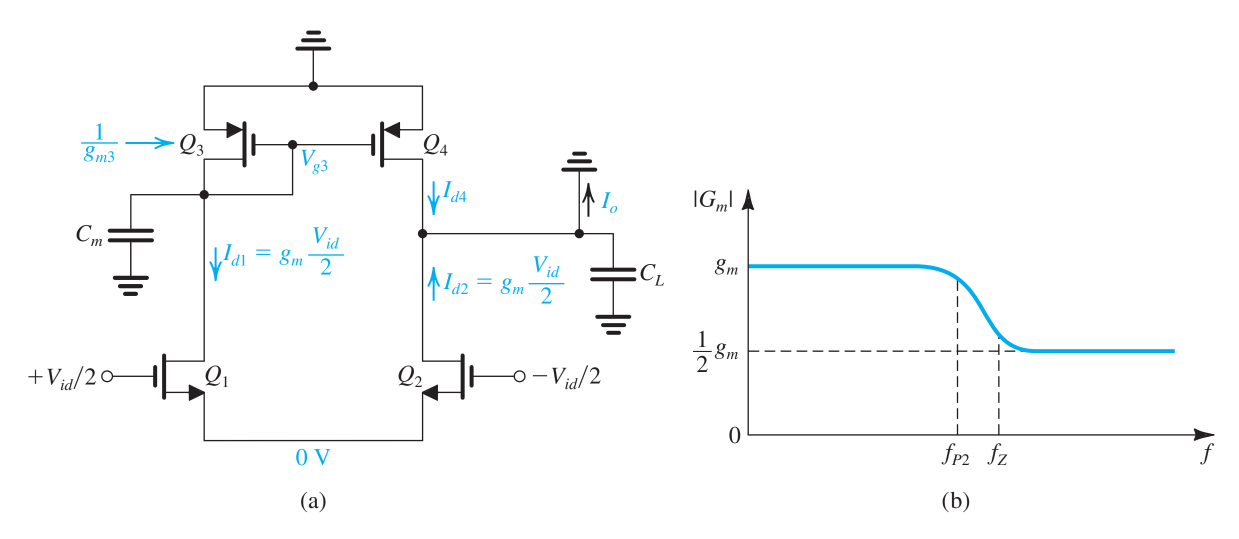


Figure 10.37 **(a)** Frequency–response analysis of the active-loaded MOS differential amplifier. **(b)** The overall transconductance G_m as a function of frequency.

and C_{gs4} but also includes C_{gd1}, C_{db1}, and C_{db3},

$$C_m = C_{gd1} + C_{db1} + C_{db3} + C_{gs3} + C_{gs4} \tag{10.133}$$

Capacitance C_L includes C_{gd2}, C_{db2}, C_{db4}, and C_{gd4} as well as an actual load capacitance and/or the input capacitance of a subsequent stage (C_x),

$$C_L = C_{gd2} + C_{db2} + C_{gd4} + C_{db4} + C_x \tag{10.134}$$

These two capacitances primarily determine the dependence of the differential gain of this amplifier on frequency.

The overall voltage gain of the differential amplifier will be determined by multiplying its short-circuit transconductance G_m by the total impedance at the output mode. As indicated in Fig. 10.37(a) the input differential signal V_{id} is applied in a balanced fashion and the output node is short-circuited to ground in order to determine the transconductance G_m; $G_m \equiv I_o/V_{id}$. Obviously, because of the output short circuit, C_L will have no effect on G_m. Transistor Q_1 will conduct a drain current signal of $g_m V_{id}/2$, which, neglecting r_{o1}, flows through the parallel combination of the diode-connected transistor Q_3 and C_m. Neglecting the resistances r_{oi} and r_{o3}, which are much larger than the resistance $(1/g_{m3})$ of Q_3 we have

$$V_{g3} = -\frac{g_m V_{id}/2}{g_{m3} + sC_m} \tag{10.135}$$

In response to V_{g3}, transistor Q_4 conducts a drain current I_{d4},

$$I_{d4} = -g_{m4} V_{g3} = \frac{g_{m4} g_m V_{id}/2}{g_{m3} + sC_m}$$

Since $g_{m3} = g_{m4}$, this equation reduces to

$$I_{d4} = \frac{g_m V_{id}/2}{1 + s\dfrac{C_m}{g_{m3}}} \tag{10.136}$$

Now, at the output node the total output current that flows through the short circuit is

$$I_o = I_{d4} + I_{d2}$$

$$= \frac{g_m V_{id}/2}{1 + s\dfrac{C_m}{g_{m3}}} + g_m V_{id}/2 \tag{10.137}$$

We can now obtain G_m as

$$G_m \equiv \frac{I_o}{V_{id}} = g_m \frac{1 + s\dfrac{C_m}{2g_{m3}}}{1 + s\dfrac{C_m}{g_{m3}}} \tag{10.138}$$

Thus, as expected, the low-frequency value of G_m is equal to g_m of Q_1 and Q_2. At high frequencies, G_m acquires a pole and a zero, the frequencies of which are

$$f_{P2} = \frac{g_{m3}}{2\pi C_m} \tag{10.139}$$

and

$$f_Z = \frac{2g_{m3}}{2\pi C_m} \tag{10.140}$$

That is, the zero frequency is twice that of the pole. Since C_m is approximately equal to $C_{gs2} + C_{gs4} = 2C_{gs}$, we also have

$$f_{P2} = \frac{g_{m3}}{2\pi C_m} \simeq \frac{g_{m3}}{2\pi (2C_{gs})} \simeq f_T/2 \tag{10.141}$$

and

$$f_Z \simeq f_T \tag{10.142}$$

where f_T is the unity-gain frequency of the MOSFET Q_3. Thus, the **mirror pole and zero** occur at very high frequencies. Nevertheless, their effect can be significant.

Figure 10.37(b) shows a sketch of the magnitude of G_m versus frequency. It is interesting and useful to observe that the path of the signal current produced by Q_1 has a transfer function different from that of the signal current produced by Q_2. It is the first signal that encounters C_m and experiences the mirror pole. This observation leads to an interesting view of the effect of C_m on the overall transconductance G_m of the differential amplifier. As we learned in Section 9.5, at low frequencies I_{d1} is replicated by the mirror Q_3–Q_4 in the drain of Q_4 as I_{d4}, which adds to I_{d2} to provide a factor-of-2 increase in G_m (thus making G_m equal to g_m, which is double the value available without the current mirror). Now, at high frequencies C_m acts as a short circuit causing V_{g3} to be zero, and hence I_{d4} will be zero, that is, destroying the action of the current mirror and reducing G_m to $g_m/2$, as borne out by the sketch in Fig. 10.37(b).

Having determined the short-circuit output current I_o, we now multiply it by the total impedance between the output node and ground to determine the output voltage V_o,

$$V_o = I_o \frac{1}{\frac{1}{R_o} + sC_L}$$

$$= G_m V_{id} \frac{R_o}{1 + sC_L R_o}$$

where

$$R_o = r_{o2} \parallel r_{o4} \tag{10.143}$$

Thus,

$$\frac{V_o}{V_{id}} = (g_m R_o)\left(\frac{1 + s\frac{C_m}{2g_{m3}}}{1 + s\frac{C_m}{g_{m3}}}\right)\left(\frac{1}{1 + sC_L R_o}\right) \tag{10.144}$$

Thus, in addition to the pole and zero of G_m, the gain of the differential amplifier will have a pole with frequency f_{P1},

$$f_{P1} = \frac{1}{2\pi C_L R_o} \tag{10.145}$$

This, of course, is entirely expected, and in fact this output pole is often dominant, especially when a large load capacitance is present.

Example 10.12

Consider an active-loaded MOS differential amplifier of the type shown in Fig. 10.37(a). Assume that for all transistors, $W/L = 7.2\ \mu\text{m}/0.36\ \mu\text{m}$, $C_{gs} = 20$ fF, $C_{gd} = 5$ fF, and $C_{db} = 5$ fF. Also, let $\mu_n C_{ox} = 387\ \mu\text{A/V}^2$, $\mu_p C_{ox} = 86\ \mu\text{A/V}^2$, $V'_{An} = 5$ V/µm, and $|V'_{Ap}| = 6$ V/µm. The bias current $I = 0.2$ mA, and the bias current source has an output resistance $R_{SS} = 25$ kΩ and an output capacitance $C_{SS} = 0.2$ pF. In addition to the capacitances introduced by the transistors at the output node, there is a capacitance C_x of 25 fF. It is required to determine the low-frequency values of A_d, A_{cm}, and CMRR. It is also required to find the poles and zero of A_d and the dominant pole of CMRR.

Solution

Since $I = 0.2$ mA, each of the four transistors is operating at a bias current of 100 µA. Thus, for Q_1 and Q_2,

$$100 = \frac{1}{2} \times 387 \times \frac{7.2}{0.36} \times V_{OV}^2$$

which leads to

$$V_{OV} = 0.16\ \text{V}$$

Example 10.12 *continued*

Thus,

$$g_m = g_{m1} = g_{m2} = \frac{2 \times 0.1}{0.16} = 1.25 \text{ mA/V}$$

$$r_{o1} = r_{o2} = \frac{5 \times 0.36}{0.1} = 18 \text{ k}\Omega$$

For Q_3 and Q_4 we have

$$100 = \frac{1}{2} \times 86 \times \frac{7.2}{0.36} V_{OV3,4}^2$$

Thus,

$$|V_{OV3,4}| = 0.34 \text{ V}$$

and

$$g_{m3} = g_{m4} = \frac{2 \times 0.1}{0.34} = 0.6 \text{ mA/V}$$

$$r_{o3} = r_{o4} = \frac{6 \times 0.36}{0.1} = 21.6 \text{ k}\Omega$$

The low-frequency value of the differential gain can be determined from

$$A_d = g_m(r_{o2} \parallel r_{o4})$$
$$= 1.25(18 \parallel 21.6) = 12.3 \text{ V/V}$$

The low-frequency value of the common-mode gain can be determined from Eq. (9.157) as

$$A_{cm} = -\frac{1}{2g_{m3}R_{SS}}$$
$$= -\frac{1}{2 \times 0.6 \times 25} = -0.033 \text{ V/V}$$

The low-frequency value of the CMRR can now be determined as

$$\text{CMRR} = \frac{|A_d|}{|A_{cm}|} = \frac{12.3}{0.033} = 369$$

or,

$$20 \log 369 = 51.3 \text{ dB}$$

To determine the poles and zero of A_d we first compute the values of the two pertinent capacitances C_m and C_L. Using Eq. (10.133),

$$C_m = C_{gd1} + C_{db1} + C_{db3} + C_{gs3} + C_{gs4}$$
$$= 5 + 5 + 5 + 20 + 20 = 55 \text{ fF}$$

Capacitance C_L is found using Eq. (10.134) as

$$C_L = C_{gd2} + C_{db2} + C_{gd4} + C_{db4} + C_x$$
$$= 5 + 5 + 5 + 5 + 25 = 45 \text{ fF}$$

Now, the poles and zero of A_d can be found from Eqs. (10.145), (10.139), and (10.140) as

$$f_{P1} = \frac{1}{2\pi C_L R_o}$$
$$= \frac{1}{2\pi \times C_L (r_{o2} \| r_{o4})}$$
$$= \frac{1}{2\pi \times 45 \times 10^{-15} (18 \| 21.6) 10^3}$$
$$= 360 \text{ MHz}$$

$$f_{P2} = \frac{g_{m3}}{2\pi C_m} = \frac{0.6 \times 10^{-3}}{2\pi \times 55 \times 10^{-15}} = 1.74 \text{ GHz}$$

$$f_Z = 2 f_{P2} = 3.5 \text{ GHz}$$

Thus the dominant pole is that produced by C_L at the output node. As expected, the pole and zero of the mirror are at much higher frequencies.

The dominant pole of the CMRR is at the location of the common-mode-gain zero introduced by C_{SS} and R_{SS}, that is,

$$f_Z = \frac{1}{2\pi C_{SS} R_{SS}}$$
$$= \frac{1}{2\pi \times 0.2 \times 10^{-12} \times 25 \times 10^3}$$
$$= 31.8 \text{ MHz}$$

Thus, the CMRR begins to decrease at 31.8 MHz, which is much lower than f_{P1}.

EXERCISE

10.28 A bipolar current-mirror-loaded differential amplifier is biased with a current source $I = 1$ mA. The transistors are specified to have $|V_A| = 100$ V. The total capacitance at the output node is 2 pF. Find the dc value, and the frequency of the dominant high-frequency pole, of the differential voltage gain.

Ans. 2000 V/V; 0.8 MHz

A

10.8 Other Wideband Amplifier Configurations

Thus far, we have studied one wideband amplifier configuration: the cascode amplifier (Section 10.5). Cascoding can, of course, be applied to differential amplifiers to obtain wideband differential amplification. In this section we discuss a number of other circuit configurations that are capable of achieving wide bandwidths.

10.8.1 Obtaining Wideband Amplification by Source and Emitter Degeneration

As we discussed in Chapter 7, adding a resistance in the source (emitter) lead of a CS (CE) amplifier can result in a number of performance improvements at the expense of a reduction in voltage gain. Extension of the amplifier bandwidth, which is the topic of interest to us in this section, is among those improvements.

Figure 10.38(a) shows a common-source amplifier with a source-degeneration resistance R_s. As indicated in Fig. 10.38(b), the output of the amplifier can be modeled at low frequencies by a controlled current-source $G_m V_i$ and an output resistance R_o, where, for the usual case of $g_m r_o \gg 1$, the transconductance G_m can be shown to be given by

$$G_m \simeq \frac{g_m}{1 + g_m R_s} \tag{10.146}$$

and the output resistance is given by Eq. (8.61), that is,

$$R_o \simeq r_o(1 + g_m R_s) \tag{10.147}$$

Thus, source degeneration reduces the transconductance and increases the output resistance by the same factor, $(1 + g_m R_s)$. The low-frequency voltage gain can be obtained as

$$A_M = \frac{V_o}{V_{\text{sig}}} = -G_m(R_o \| R_L) = -G_m R_L' \tag{10.148}$$

where

$$R_L' = R_L \| R_o \tag{10.149}$$

Let's now consider the high-frequency response of the source-degenerated amplifier. Figure 10.38(c) shows the amplifier, indicating the capacitances C_{gs} and C_{gd}. A capacitance C_L that *includes* the MOSFET capacitance C_{db} is also shown at the output. The method of open-circuit time constants can be employed to obtain an estimate of the 3-dB frequency f_H. Toward that end, we show in Fig. 10.38(d) the circuit for determining R_{gd}, which is the resistance seen by C_{gd}. Straightforward analysis yields

$$R_{gd} = R_{\text{sig}}(1 + G_m R_L') + R_L' \tag{10.150}$$

Note that the expression for R_{gd} in Eq. (10.150) is similar to that for the CS amplifier in Eq. (10.85) with g_m replaced with G_m, and $R_L' = r_o \| R_L$ of the CS replaced with $R_L' = R_o \| R_L$ for the source-degenerated case.

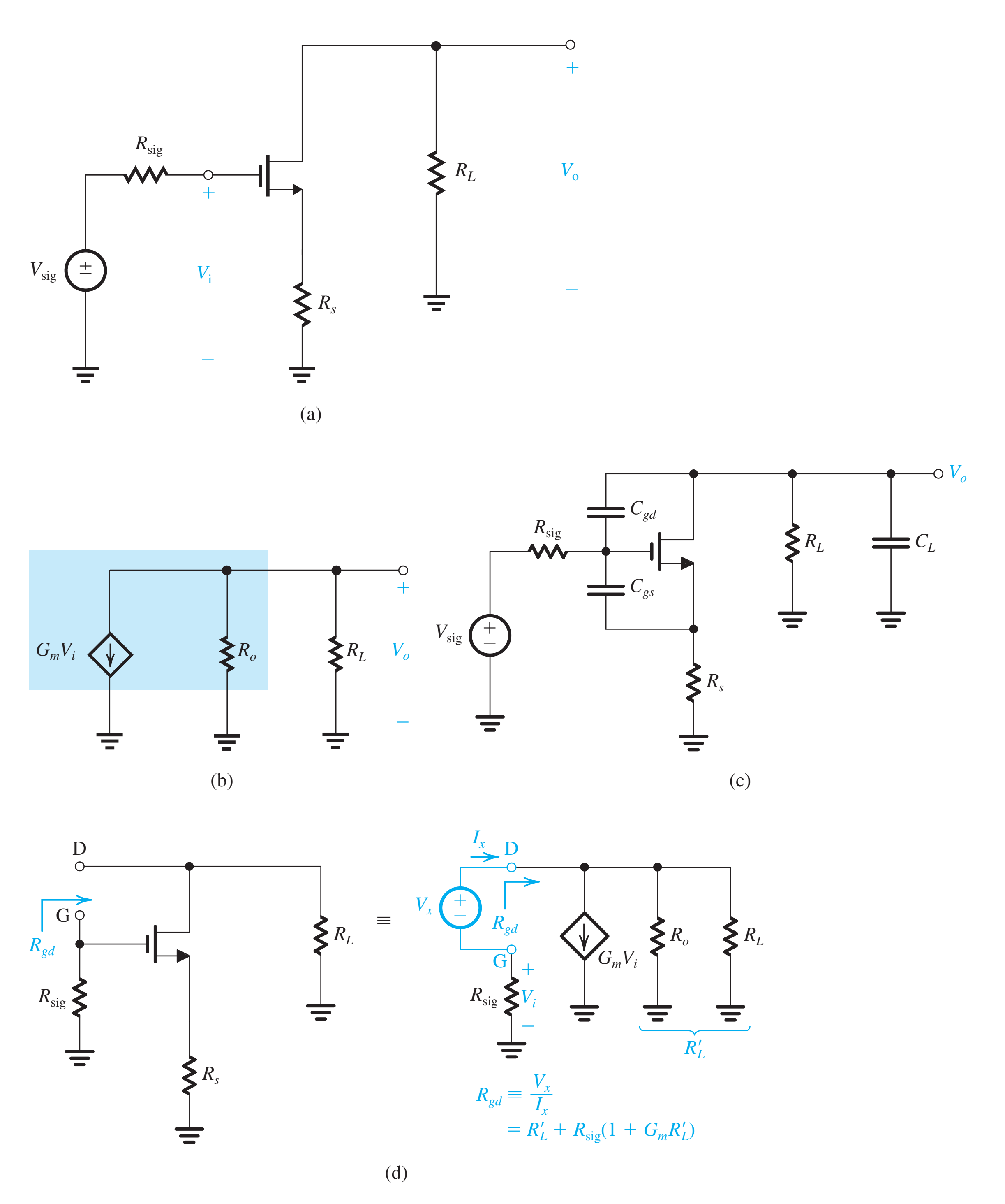


Figure 10.38 **(a)** The CS amplifier circuit, with a source resistance R_s. **(b)** Equivalent-circuit representation of the amplifier output. **(c)** The circuit prepared for frequency-response analysis. **(d)** Determining the resistance R_{gd} seen by the capacitance C_{gd}.

The formula for R_{C_L} can be seen to be simply

$$R_{C_L} = R_L \parallel R_o = R'_L \tag{10.151}$$

The formula for R_{gs} is the most difficult to derive, and the derivation should be performed with the hybrid-π model explicitly utilized. Straightforward, though somewhat tedious, circuit analysis yields (for $g_m r_o \gg 1$),

$$R_{gs} \simeq \frac{R_{sig} + R_s + R_{sig}R_s/(r_o + R_L)}{1 + g_m R_s\left(\dfrac{r_o}{r_o + R_L}\right)} \tag{10.152}$$

Next we compute τ_H,

$$\tau_H = C_{gs}R_{gs} + C_{gd}R_{gd} + C_L R_{C_L} \tag{10.153}$$

and use it to determine f_H,

$$f_H = \frac{1}{2\pi\tau_H}$$

It is interesting and instructive to consider the case when R_{sig} is relatively large: The frequency response will be dominated by the Miller multiplication of C_{gd}. Another way for saying this is that $C_{gd}R_{gd}$ will be the largest of the three open-circuit time constants that make up τ_H in Eq. (10.153), enabling us to approximate τ_H as

$$\tau_H \simeq C_{gd}R_{gd} \tag{10.154}$$

and correspondingly to obtain f_H as

$$f_H \simeq \frac{1}{2\pi C_{gd}R_{gd}} \tag{10.155}$$

Now, as R_s is increased, the gain magnitude, $|A_M| = G_m R'_L$, will decrease, causing R_{gd} to decrease (Eq. 10.150), which in turn causes f_H to increase (Eq. 10.155). To highlight the trade-off between gain and bandwidth that R_s affords the designer, let us simplify the expression for R_{gd} in Eq. (10.150) by assuming that $G_m R'_L \gg 1$ and $G_m R_{sig} \gg 1$, thus

$$R_{gd} \simeq G_m R'_L R_{sig} = |A_M| R_{sig}$$

which can be substituted in Eq. (10.155) to obtain

$$f_H = \frac{1}{2\pi C_{gd}R_{sig}|A_M|} \tag{10.156}$$

which very clearly shows the gain–bandwidth trade-off. The gain–bandwidth product remains constant at

$$\text{Gain–bandwidth product} = |A_M| f_H = \frac{1}{2\pi C_{gd}R_{sig}} \tag{10.157}$$

In practice, however, the other capacitances will play a role in determining f_H, and the gain–bandwidth product will decrease somewhat as R_s is increased.

EXERCISE

10.29 Consider a CS amplifier having $g_m = 2$ mA/V, $r_o = 20$ kΩ, $R_L = 20$ kΩ, $R_{sig} = 20$ kΩ, $C_{gs} = 20$ fF, $C_{gd} = 5$ fF, and $C_L = 5$ fF. (a) Find the voltage gain A_M and the 3-dB frequency f_H (using the method of open-circuit time constants) and hence the gain–bandwidth product. (b) Repeat (a) for the case in which a resistance R_s is connected in series with the source terminal with a value selected so that $g_m R_s = 2$.

Ans. (a) –20 V/V, 61.2 MHz, 1.22 GHz; (b) –10 V/V, 109 MHz, 1.1 GHz

10.8.2 The CD–CS, CC–CE, and CD–CE Configurations

In Section 8.7.1 we discussed the performance improvements obtained by preceding the CS and CE amplifiers by a buffer implemented by a CD or a CC amplifier, as in the circuits shown in Fig. 10.39. A major advantage of each of these circuits is wider bandwidth than that obtained in the CS or CE stage alone. To see how this comes about, consider as an example the CD–CS amplifier in Fig 10.39(a) and note that the CS transistor Q_2 will still exhibit a Miller effect that results in a large input capacitance, C_{in2}, between its gate and ground. However, the resistance that this capacitance interacts with will be much lower than R_{sig}; the buffering action of the source follower causes a relatively low resistance, approximately equal to $1/g_{m1}$, to appear between the source of Q_1 and ground across C_{in2}.

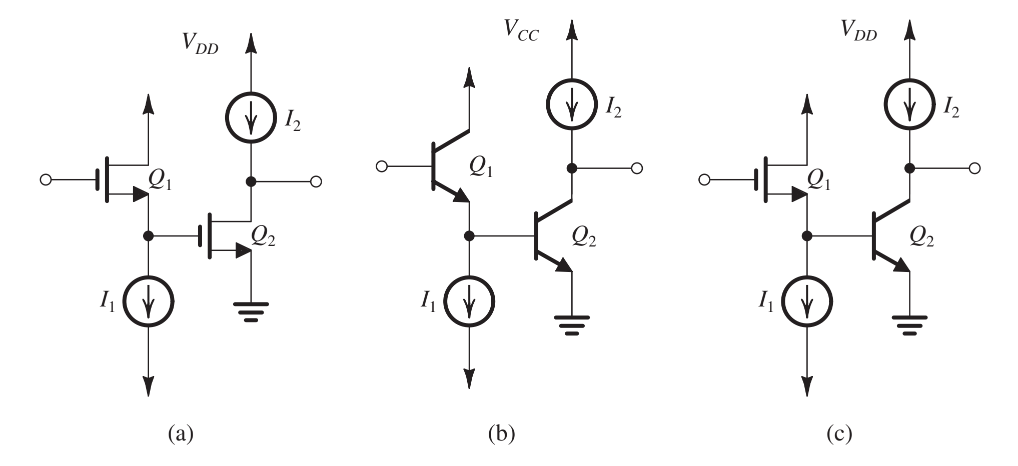


Figure 10.39 **(a)** CD–CS amplifier. **(b)** CC–CE amplifier. **(c)** CD–CE amplifier.

Example 10.13

Consider a CC–CE amplifier such as that in Fig. 10.39(b) with the following specifications: $I_1 = I_2 = 1$ mA and identical transistors with $\beta = 100$, $f_T = 400$ MHz, and $C_\mu = 2$ pF. Let the amplifier be fed with a source V_{sig} having a resistance $R_{sig} = 4$ kΩ, and assume a load resistance of 4 kΩ. Find the voltage gain A_M,

Example 10.13 *continued*

and estimate the 3-dB frequency, f_H. Compare the results with those obtained with a CE amplifier operating under the same conditions. For simplicity, neglect r_o and r_x.

Solution

At an emitter bias current of 1 mA, Q_1 and Q_2 have

$$g_m = 40 \text{ mA/V}$$
$$r_e = 25 \ \Omega$$
$$r_\pi = \frac{\beta}{g_m} = \frac{100}{40} = 2.5 \text{ k}\Omega$$
$$C_\pi + C_\mu = \frac{g_m}{\omega_T} = \frac{g_m}{2\pi f_T}$$
$$= \frac{40 \times 10^{-3}}{2\pi \times 400 \times 10^6} = 15.9 \text{ pF}$$
$$C_\mu = 2 \text{ pF}$$
$$C_\pi = 13.9 \text{ pF}$$

The voltage gain A_M can be determined from the circuit shown in Fig. 10.40(a) as follows:

$$R_{\text{in}2} = r_{\pi 2} = 2.5 \text{ k}\Omega$$
$$R_{\text{in}} = (\beta_1 + 1)(r_{e1} + R_{\text{in}2})$$
$$= 101(0.025 + 2.5) = 255 \text{ k}\Omega$$
$$\frac{V_{b1}}{V_{\text{sig}}} = \frac{R_{\text{in}}}{R_{\text{in}} + R_{\text{sig}}} = \frac{255}{255 + 4} = 0.98 \text{ V/V}$$
$$\frac{V_{b2}}{V_{b1}} = \frac{R_{\text{in}2}}{R_{\text{in}2} + r_{e1}} = \frac{2.5}{2.5 + 0.025} = 0.99 \text{ V/V}$$
$$\frac{V_o}{V_{b2}} = -g_{m2}R_L = -40 \times 4 = -160 \text{ V/V}$$

Thus,

$$A_M = \frac{V_o}{V_{\text{sig}}} = -160 \times 0.99 \times 0.98 = -155 \text{ V/V}$$

To determine f_H we use the method of open-circuit time constants. Figure 10.40(b) shows the circuit with V_{sig} set to zero and the four capacitances indicated. Capacitance $C_{\mu 1}$ sees a resistance $R_{\mu 1}$,

$$R_{\mu 1} = R_{\text{sig}} \| R_{\text{in}}$$
$$= 4 \| 255 = 3.94 \text{ k}\Omega$$

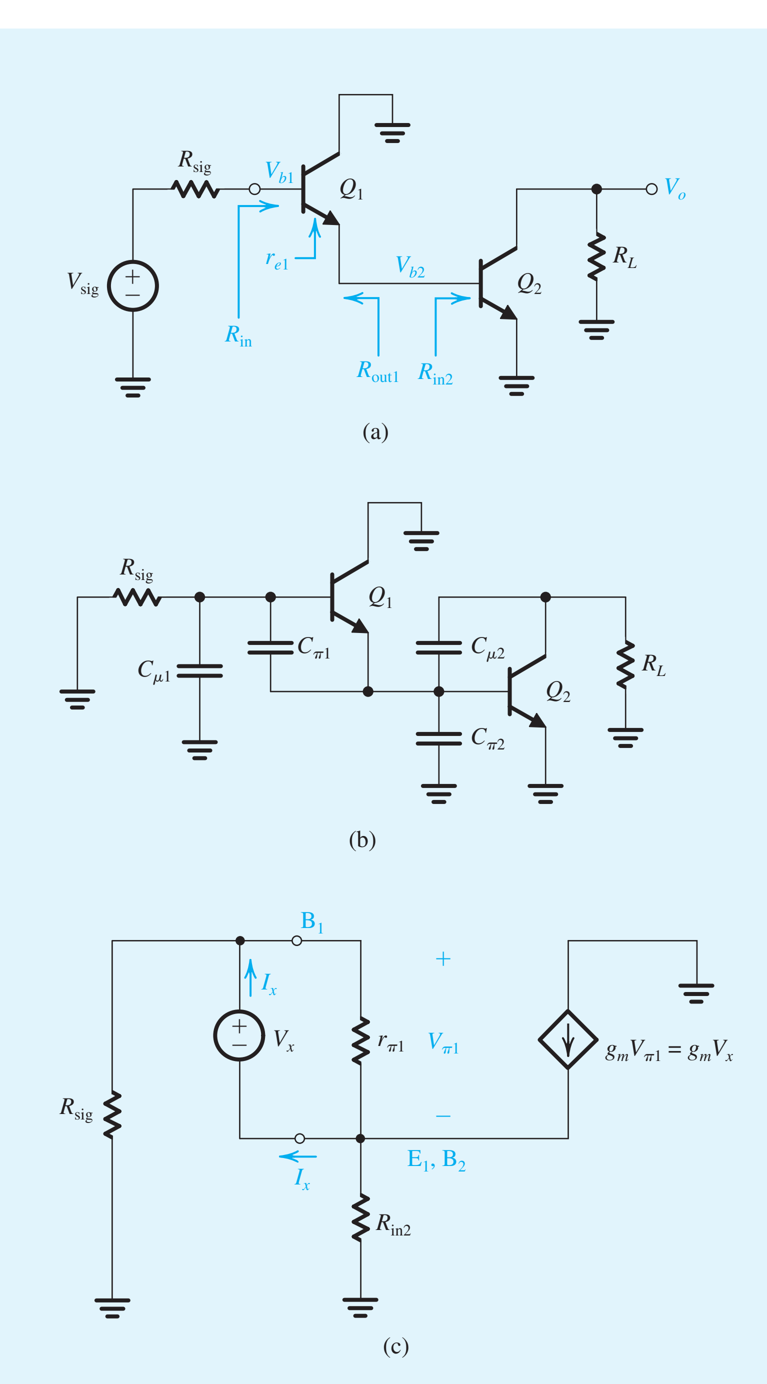


Figure 10.40 Circuits for Example 10.14: **(a)** the CC–CE circuit prepared for low-frequency, small-signal analysis; **(b)** the circuit at high frequencies, with V_{sig} set to zero to enable determination of the open-circuit time constants; **(c)** equivalent circuit for the determination of $R_{\pi 1}$; **(d)** a CE amplifier for comparison.

Example 10.13 *continued*

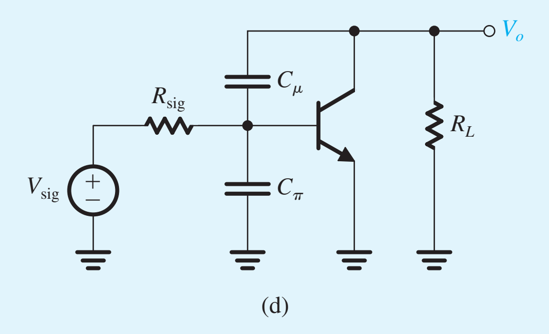


Figure 10.40 *continued*

To find the resistance $R_{\pi 1}$ seen by capacitance $C_{\pi 1}$ we refer to the equivalent circuit in Fig. 10.40(c). Analysis of this circuit results in

$$R_{\pi 1} \equiv \frac{V_x}{I_x} = \frac{R_{sig} + R_{in2}}{1 + \dfrac{R_{sig}}{r_{\pi 1}} + \dfrac{R_{in2}}{r_{e1}}}$$

$$= \frac{4000 + 2500}{1 + \dfrac{4000}{2500} + \dfrac{2500}{25}} = 63.4\ \Omega$$

Capacitance $C_{\pi 2}$ sees a resistance $R_{\pi 2}$,

$$R_{\pi 2} = R_{in2} \parallel R_{out1}$$

$$= r_{\pi 2} \parallel \left[r_{e1} + \frac{R_{sig}}{\beta_1 + 1} \right]$$

$$= 2500 \parallel \left[25 + \frac{4000}{101} \right] = 63\ \Omega$$

Capacitance $C_{\mu 2}$ sees a resistance $R_{\mu 2}$. To determine $R_{\mu 2}$ we refer to the analysis of the frequency response of the CE amplifier in Section 10.4.4 to obtain

$$R_{\mu 2} = (1 + g_{m2} R_L)(R_{in2} \parallel R_{out1}) + R_L$$

$$= (1 + 40 \times 4)\left[2500 \parallel \left(25 + \frac{4000}{101} \right) \right] + 4000$$

$$= 14{,}143\ \Omega \simeq 14.1\ \text{k}\Omega$$

We now can determine τ_H from

$$\tau_H = C_{\mu 1}R_{\mu 1} + C_{\pi 1}R_{\pi 1} + C_{\mu 2}R_{\mu 2} + C_{\pi 2}R_{\pi 2}$$
$$= 2 \times 3.94 + 13.9 \times 0.0634 + 2 \times 14.1 + 13.9 \times 0.063$$
$$= 7.88 + 0.88 + 28.2 + 0.88 = 37.8 \text{ ns}$$

We observe that $C_{\pi 1}$ and $C_{\pi 2}$ play minor roles in determining the high-frequency response. As expected, $C_{\mu 2}$ through the Miller effect plays the most significant role. Capacitor $C_{\mu 1}$, which interacts directly with $(R_{\text{sig}} \| R_{\text{in}})$, also plays an important role. The 3-dB frequency f_H can be found as follows:

$$f_H = \frac{1}{2\pi\tau_H} = \frac{1}{2\pi \times 37.8 \times 10^{-9}} = 4.2 \text{ MHz}$$

For comparison, we evaluate A_M and f_H of a CE amplifier operating under the same conditions. Refer to Fig. 10.40(d). The voltage gain A_M is given by

$$A_M = \frac{R_{\text{in}}}{R_{\text{in}} + R_{\text{sig}}}(-g_m R_L)$$
$$= \frac{r_\pi}{r_\pi + R_{\text{sig}}}(-g_m R_L)$$
$$= \frac{2.5}{2.5 + 4}(-40 \times 4)$$
$$= -61.5 \text{ V/V}$$
$$R_\pi = r_\pi \| R_{\text{sig}} = 2.5 \| 4 = 1.54 \text{ k}\Omega$$
$$R_\mu = (1 + g_m R_L)(R_{\text{sig}} \| r_\pi) + R_L$$
$$= (1 + 40 \times 4)(4 \| 2.5) + 4$$
$$= 251.7 \text{ k}\Omega$$

Thus,

$$\tau_H = C_\pi R_\pi + C_\mu R_\mu$$
$$= 13.9 \times 1.54 + 2 \times 251.7$$
$$= 21.4 + 503.4 = 524.8 \text{ ns}$$

Observe the dominant role played by C_μ. The 3-dB frequency f_H is

$$f_H = \frac{1}{2\pi\tau_H} = \frac{1}{2\pi \times 524.8 \times 10^{-9}} = 303 \text{ kHz}$$

Thus, including the buffering transistor Q_1 increases the gain, $|A_M|$, from 61.5 V/V to 155 V/V—a factor of 2.5—and increases the bandwidth from 303 kHz to 4.2 MHz—a factor of 13.9! The gain–bandwidth product is increased from 18.63 MHz to 651 MHz—a factor of 35!

10.8.3 The CC–CB and CD–CG Configurations

In Section 8.7.3 we showed that preceding a CB or CG transistor with a buffer implemented with a CC or a CD transistor solves the low-input-resistance problem of the CB and CG amplifiers. Examples of the resulting compound-transistor amplifiers are shown in Fig. 10.41. Since in each of these circuits, neither of the two transistors suffers from the Miller effect, the resulting amplifiers have even wider bandwidths than those achieved in the compound amplifier stages of the last section. To illustrate, consider as an example the circuit in Fig. 10.41(a).[10] The low-frequency analysis of this circuit in Section 8.7.3 provides for the input resistance,

$$R_{in} = (\beta_1 + 1)(r_{e1} + r_{e2}) \tag{10.158}$$

which for $r_{e1} = r_{e2} = r_e$ and $\beta_1 = \beta_2 = \beta$ becomes

$$R_{in} = 2r_\pi \tag{10.159}$$

If a load resistance R_L is connected at the output, the voltage gain V_o/V_i will be

$$\frac{V_o}{V_i} = \frac{\alpha_2 R_L}{r_{e1} + r_{e2}} = \frac{1}{2} g_m R_L \tag{10.160}$$

Now, if the amplifier is fed with a voltage signal V_{sig} from a source with a resistance R_{sig}, the overall voltage gain will be

$$\frac{V_o}{V_{sig}} = \frac{1}{2}\left(\frac{R_{in}}{R_{in} + R_{sig}}\right)(g_m R_L) \tag{10.161}$$

The high-frequency analysis is illustrated in Fig. 10.42(a). Here we have drawn the hybrid-π equivalent circuit for each of Q_1 and Q_2. Recalling that the two transistors are operating at equal bias currents, their corresponding model components will be equal (i.e., $r_{\pi 1} = r_{\pi 2}$, $C_{\pi 1} = C_{\pi 2}$, etc.). With this in mind the reader should be able to see that $V_{\pi 1} = -V_{\pi 2}$ and the horizontal line through the node labeled E in Fig. 10.42(a) can be deleted. Thus the circuit reduces to that in Fig. 10.42(b). This is a very attractive outcome because the circuit shows clearly the two poles that determine the high-frequency response: The pole at the input, with a frequency f_{P1}, is

$$f_{P1} = \frac{1}{2\pi\left(\dfrac{C_\pi}{2} + C_\mu\right)(R_{sig} \| 2r_\pi)} \tag{10.162}$$

and the pole at the output, with a frequency f_{P2}, is

$$f_{P2} = \frac{1}{2\pi C_\mu R_L} \tag{10.163}$$

This result is also intuitively obvious: The input impedance at B_1 of the circuit in Fig. 10.42(a) consists of the series connection of $r_{\pi 1}$ and $r_{\pi 2}$ in parallel with the series connection of $C_{\pi 1}$ and $C_{\pi 2}$. Then there is $C_{\mu 1}$ in parallel. At the output, we simply have R_L in parallel with C_μ.

[10]The results derived for the circuit in Fig. 10.41(a) apply directly to the circuit of Fig. 10.41(b) and with appropriate change of variables to the MOS circuit of Fig. 10.41(c).

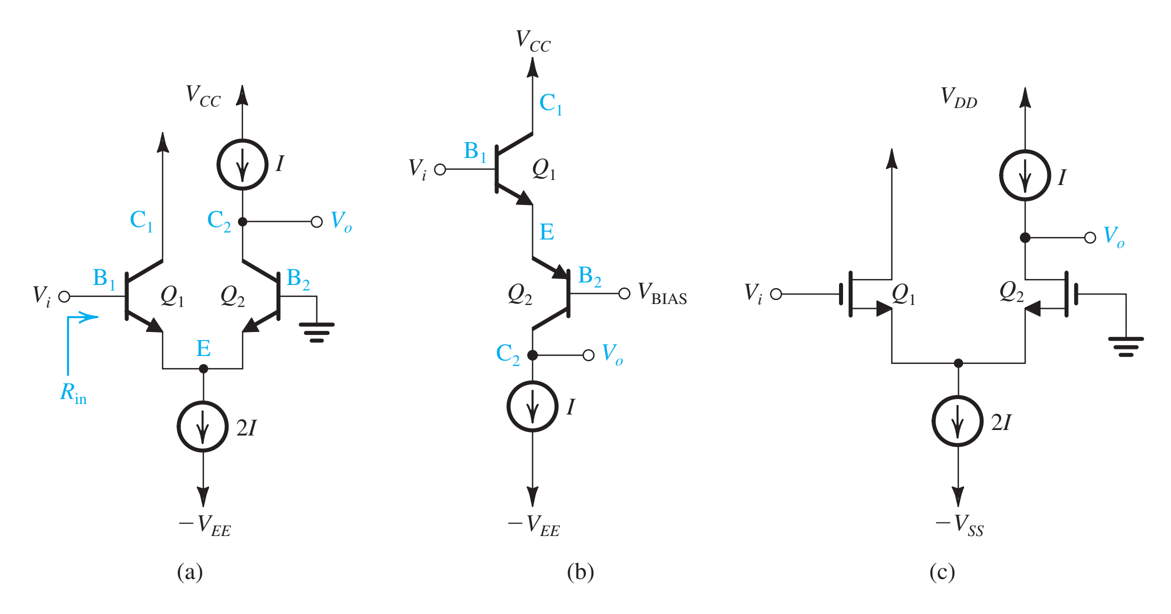


Figure 10.41 (**a**) A CC–CB amplifier. (**b**) Another version of the CC–CB circuit with Q_2 implemented using a *pnp* transistor. (**c**) The MOSFET version of the circuit in (**a**).

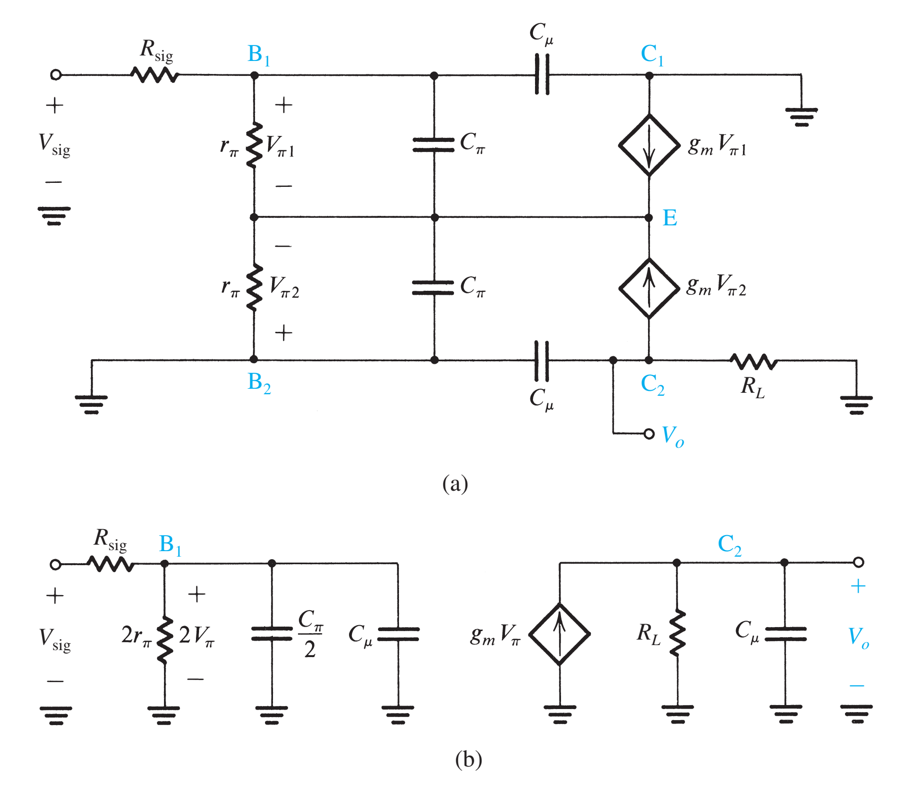


Figure 10.42 (**a**) Equivalent circuit for the amplifier in Fig. 10.41(a). (**b**) Simplified equivalent circuit. Note that the equivalent circuits in (**a**) and (**b**) also apply to the circuit shown in Fig. 10.41(b). In addition, they can be easily adapted for the MOSFET circuit in Fig. 10.41(c), with $2r_\pi$ eliminated, C_π replaced with C_{gs}, C_μ replaced with C_{gd}, and V_π replaced with V_{gs}.

Whether one of the two poles is dominant will depend on the relative values of R_{sig} and R_L. If the two poles are close to each other, then the 3-dB frequency f_H can be determined either by exact analysis—that is, finding the frequency at which the gain is down by 3 dB—or by using the approximate formula in Eq. (10.77),

$$f_H \simeq 1 \Big/ \sqrt{\frac{1}{f_{P1}^2} + \frac{1}{f_{P2}^2}} \tag{10.164}$$

EXERCISE

10.30 For the CC–CB amplifier of Fig. 10.41(a), let $I = 0.5$ mA, $\beta = 100$, $C_\pi = 6$ pF, $C_\mu = 2$ pF, $R_{sig} = 10$ kΩ, and $R_L = 10$ kΩ. Find the low-frequency overall voltage gain A_M, the frequencies of the poles, and the 3-dB frequency f_H. Find f_H both exactly and using the approximate formula in Eq. (10.164).
Ans. 50 V/V; 6.4 MHz and 8 MHz; f_H by exact evaluation = 4.6 MHz; f_H using Eq. (10.164) = 5 MHz.

Summary

- The coupling and bypass capacitors utilized in discrete-circuit amplifiers cause the amplifier gain to fall off at low frequencies. In the CS amplifier, the capacitors do not interact, and the frequencies of the low-frequency poles can be estimated by considering each of these capacitors separately and determining the resistance seen by the capacitor. The highest-frequency pole is the one that determines the lower 3-dB frequency f_L. In the CE amplifier, the capacitors interact, and thus the poles cannot be easily determined. Rather the method of short-circuit time constants can be used to obtain an estimate of the 3-dB frequency, f_L.

- Both the MOSFET and the BJT have internal capacitive effects that can be modeled by augmenting the device hybrid-π model with capacitances. Usually at least two capacitances are needed: C_{gs} and C_{gd} (C_π and C_μ for the BJT). A figure of merit for the high-frequency operation of the transistor is the frequency f_T at which the short-circuit current gain of the CS (CE) transistor reduces to unity. For the MOSFET, $f_T = g_m/2\pi\left(C_{gs} + C_{gd}\right)$, and for the BJT, $f_T = g_m/2\pi\left(C_\pi + C_\mu\right)$.

- The internal capacitances of the MOSFET and the BJT cause the amplifier gain to fall off at high frequencies. An estimate of the amplifier bandwidth is provided by the frequency f_H at which the gain drops 3 dB below its value at midband, A_M. A figure of merit for the amplifier is the gain–bandwidth product $GB = A_M f_H$. Usually, it is possible to trade off gain for increased bandwidth, with GB remaining nearly constant. For amplifiers with a dominant pole with frequency f_H, the gain falls off at a uniform 6-dB/octave (20-dB/decade) rate, reaching 0 dB at $f_t = GB$.

- The high-frequency response of the CS and CE amplifiers is severely limited by the Miller effect: The small capacitance C_{gd} $\left(C_\mu\right)$ is multiplied by a factor approximately equal to the gain from gate to drain (base to collector) $g_m R_L'$ and thus gives rise to a large capacitance at the amplifier input. The increased C_{in} interacts with the effective signal-source resistance R_{sig}' and causes the amplifier gain to have a 3-dB frequency $f_H = 1/2\pi R_{sig}' C_{in}$.

- The method of open-circuit time constants provides a simple and powerful way to obtain a reasonably good estimate of the upper 3-dB frequency f_H. The capacitors that limit the high-frequency response are considered one at time with $V_{sig} = 0$ and all the other capacitances set to zero (open circuited). The resistance seen by each capacitance is determined, and the overall time constant τ_H is obtained by summing the individual time constants. Then f_H is found as $1/2\pi\tau_H$.

- The CG and CB amplifiers do *not* suffer from the Miller effect. Thus the cascode amplifier, which consists of a cascade of CS and CG stages (CE and CB stages), can be designed to obtain wider bandwidth than that achieved

in the CS (CE) amplifier alone. The key, however, is to design the cascode so that the gain obtained in the CS (CE) stage is minimized.

- The source and emitter followers can have complex poles. Thus, their frequency response is evaluated using the complete transfer function. Followers of both types exhibit wide bandwidths.
- The high-frequency response of the differential amplifier can be obtained by considering the differential and common-mode half-circuits. The CMRR falls off at a relatively low frequency determined by the output impedance of the bias current source.
- The high-frequency response of the current-mirror-loaded differential amplifier is complicated by the fact that there are two signal paths between input and output: a direct path and one through the current mirror.
- Combining two transistors in a way that eliminates or minimizes the Miller effect can result in a much wider bandwidth. Some such configurations are presented in Section 10.8.

PROBLEMS

Computer Simulation Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as gain–bandwidth trade-off. Instructions to assist in setting up PSpice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 10.1: Low-Frequency Response of Discrete-Circuit Common-Source and Common-Emitter Amplifiers

D 10.1 For the amplifier in Fig. 10.3(a), if $R_{G1} = 2\text{ M}\Omega$, $R_{G2} = 1\text{ M}\Omega$, and $R_{sig} = 200\text{ k}\Omega$, find the value of the coupling capacitor C_{C1} (specified to one significant digit) that places the associated pole at 10 Hz or lower.

D 10.2 For the amplifier in Fig. 10.3(a), if $R_D = 10\text{ k}\Omega$, $R_L = 10\text{ k}\Omega$, and r_o is very large, find the value of C_{C2} (specified to one significant digit) that places the associated pole at 10 Hz or lower.

D 10.3 The amplifier in Fig. 10.3(a) is biased to operate at $g_m = 5\text{ mA/V}$, and $R_S = 1.8\text{ k}\Omega$. Find the value of C_S (specified to one significant digit) that places its associated pole at 100 Hz or lower. What are the actual frequencies of the pole and zero realized?

10.4 The amplifier in Fig. 10.3(a) is biased to operate at $g_m = 5\text{ mA/V}$, and has the following component values: $R_{sig} = 100\text{ k}\Omega$, $R_{G1} = 47\text{ M}\Omega$, $R_{G2} = 10\text{ M}\Omega$, $C_{C1} = 0.01\ \mu\text{F}$, $R_S = 2\text{ k}\Omega$, $C_S = 10\ \mu\text{F}$, $R_D = 4.7\text{ k}\Omega$, $R_L = 10\text{ k}\Omega$, and $C_{C2} = 1\ \mu\text{F}$. Find A_M, f_{P1}, f_{P2}, f_Z, f_{P3}, and f_L.

D 10.5 The amplifier in Fig. P10.5 is biased to operate at $g_m = 2\text{ mA/V}$. Neglect r_o.

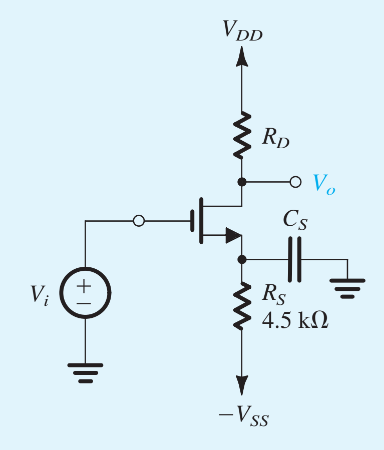


Figure P10.5

(a) Determine the value of R_D that results in a midband gain of −20 V/V.
(b) Determine the value of C_S that results in a pole frequency of 100 Hz.
(c) What is the frequency of the transmission zero introduced by C_S?
(d) Give an approximate value for the 3-dB frequency f_L.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

(e) Sketch a Bode plot for the gain of this amplifier. What does the plot tell you about the gain at dc? Does this make sense? Why or why not?

D 10.6 Figure P10.6 shows a CS amplifier biased by a constant-current source I. Let $R_{sig} = 0.5\ \text{M}\Omega$, $R_G = 2\ \text{M}\Omega$, $g_m = 3\ \text{mA/V}$, $R_D = 20\ \text{k}\Omega$, and $R_L = 10\ \text{k}\Omega$. Find A_M. Also, design the coupling and bypass capacitors to locate the three low-frequency poles at 100 Hz, 10 Hz, and 1 Hz. Use a minimum total capacitance, with the capacitors specified only to a single significant digit. What value of f_L results?

D 10.7 Figure P10.7 shows a current-biased CE amplifier operating at 100 μA from ±3-V power supplies. It employs

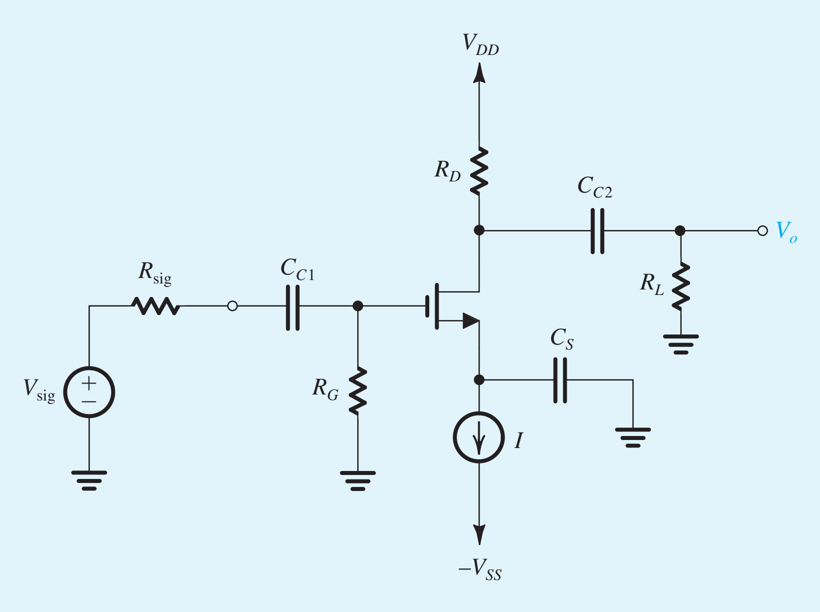


Figure P10.6

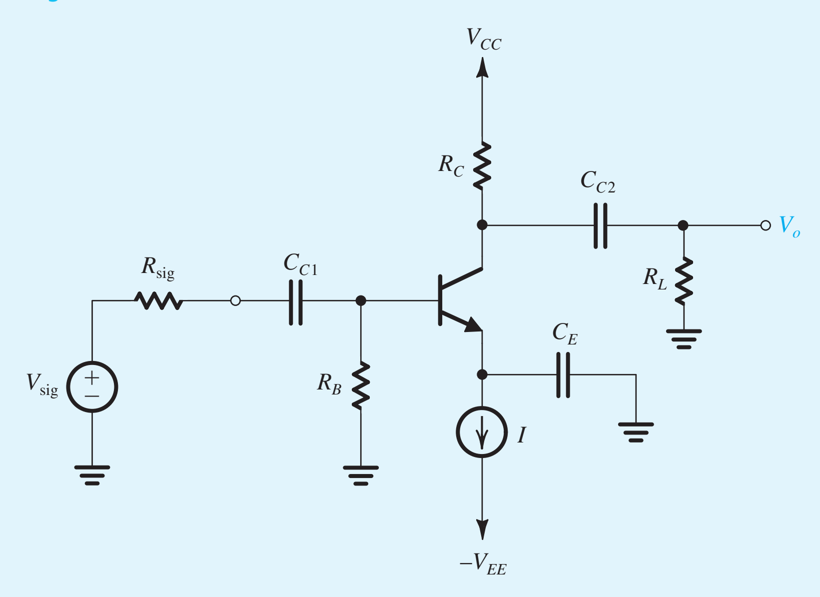


Figure P10.7

$R_C = 20$ kΩ, $R_B = 200$ kΩ, and operates between a 20-kΩ source and a 10-kΩ load. The transistor $\beta = 100$. Select C_E first, for a minimum value specified to one significant digit and providing up to 80% of f_L where f_L is to be 100 Hz. Then choose C_{C1} and C_{C2}, each specified to one significant digit, and each contributing about 10% of f_L. What f_L results? What total capacitance is needed?

10.8 Consider the common-emitter amplifier of Fig. 10.9(a) under the following conditions: $R_{sig} = 5$ kΩ, $R_{B1} = 33$ kΩ, $R_{B2} = 22$ kΩ, $R_E = 3.9$ kΩ, $R_C = 4.7$ kΩ, $R_L = 5.6$ kΩ, $V_{CC} = 5$ V. The dc emitter current can be shown to be $I_E \simeq 0.3$ mA, at which $\beta = 120$. Find the input resistance R_{in} and the midband gain A_M. If $C_{C1} = C_{C2} = 1$ µF and $C_E = 20$ µF, find the three short-circuit time constants and an estimate for f_L.

D 10.9 For the amplifier described in Problem 10.8, design the coupling and bypass capacitors for a lower 3-dB frequency of 50 Hz. Design so that the contribution of each of C_{C1} and C_{C2} to determining f_L is only 10%.

10.10 Consider the circuit of Fig. 10.9(a). For $R_{sig} = 5$ kΩ, $R_B \equiv R_{B1} \| R_{B2} = 10$ kΩ, $r_\pi = 1$ kΩ, $\beta_0 = 100$, and $R_E = 1.5$ kΩ, what is the ratio C_E/C_{C1} that makes their contributions to the determination of f_L equal?

D *10.11 For the common-emitter amplifier of Fig. P10.11, neglect r_o and assume the current source to be ideal.

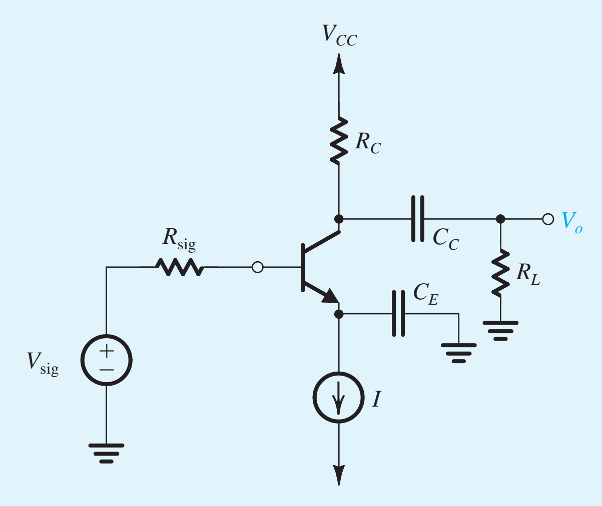


Figure P10.11

(a) Derive an expression for the midband gain.

(b) Convince yourself that the two poles caused by C_E and C_C do not interact. Find expressions for their frequencies, ω_{PE} and ω_{PC}.

(c) Give an expression for the amplifier voltage gain $V_o(s)/V_{sig}(s)$ in terms of A_M, ω_{PE}, and ω_{PC}.

(d) For $R_{sig} = R_C = R_L = 10$ kΩ, $\beta = 100$, and $I = 1$ mA, find the value of the midband gain.

(e) Select values for C_E and C_C to place the two pole frequencies a decade apart and to obtain a lower 3-dB frequency of 100 Hz while minimizing the total capacitance.

(f) Sketch a Bode plot for the gain magnitude, and estimate the frequency at which the gain becomes unity.

***10.12** The BJT common-emitter amplifier of Fig. P10.12 includes an emitter-degeneration resistance R_e.

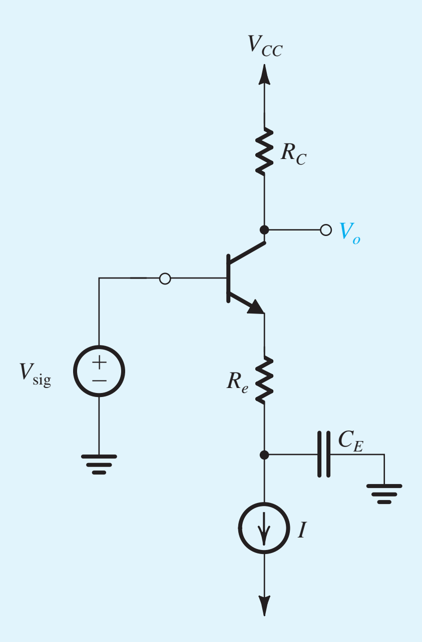


Figure P10.12

(a) Assuming $\alpha \simeq 1$, neglecting r_o, and assuming the current source to be ideal, derive an expression for the small-signal voltage gain $A(s) \equiv V_o/V_{sig}$ that applies in the midband and the low-frequency band. Hence find the midband gain A_M and the lower 3-dB frequency f_L.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

(b) Show that including R_e reduces the magnitude of A_M by a certain factor. What is this factor?

(c) Show that including R_e reduces f_L by the same factor as in (b) and thus one can use R_e to trade off gain for bandwidth.

(d) For $I = 0.25$ mA, $R_C = 10\text{ k}\Omega$, and $C_E = 10\ \mu\text{F}$, find $|A_M|$ and f_L with $R_e = 0$. Now find the value of R_e that lowers f_L by a factor of 10. What will the gain become? Sketch on the same diagram a Bode plot for the gain magnitude for both cases.

Section 10.2: Internal Capacitive Effects and the High-Frequency Model of the MOSFET and the BJT

10.13 Refer to the MOSFET high-frequency model in Fig. 10.12(a). Evaluate the model parameters for an NMOS transistor operating at $I_D = 200\ \mu\text{A}$, $V_{SB} = 1$ V, and $V_{DS} = 1.5$ V. The MOSFET has $W = 20\ \mu\text{m}$, $L = 1\ \mu\text{m}$, $t_{ox} = 8$ nm, $\mu_n = 450\text{ cm}^2/\text{V}\cdot\text{s}$, $\gamma = 0.5\text{ V}^{1/2}$, $2\phi_f = 0.65$ V, $\lambda = 0.05\text{ V}^{-1}$, $V_0 = 0.7$ V, $C_{sb0} = C_{db0} = 20$ fF, and $L_{ov} = 0.05\ \mu\text{m}$. [Recall that $g_{mb} = \chi g_m$, where $\chi = \gamma/\left(2\sqrt{2\phi_f + V_{SB}}\right)$, and that $\epsilon_{ox} = 3.45 \times 10^{-11}$ F/m.]

10.14 Find f_T for a MOSFET operating at $I_D = 200\ \mu\text{A}$ and $V_{OV} = 0.3$ V. The MOSFET has $C_{gs} = 25$ fF and $C_{gd} = 5$ fF.

10.15 Starting from the expression of f_T for a MOSFET,

$$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$$

and making the approximation that $C_{gs} \gg C_{gd}$ and that the overlap component of C_{gs} is negligibly small, show that

$$f_T \simeq \frac{1.5}{\pi L}\sqrt{\frac{\mu_n I_D}{2C_{ox}WL}}$$

Thus note that to obtain a high f_T from a given device, it must be operated at a high current. Also note that faster operation is obtained from smaller devices.

10.16 Starting from the expression for the MOSFET unity-gain frequency,

$$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$$

and making the approximation that $C_{gs} \gg C_{gd}$ and that the overlap component of C_{gs} is negligibly small, show that for an n-channel device

$$f_T \simeq \frac{3\mu_n V_{OV}}{4\pi L^2}$$

Observe that for a given channel length, f_T can be increased by operating the MOSFET at a higher overdrive voltage. Evaluate f_T for devices with $L = 0.5\ \mu\text{m}$ operated at overdrive voltages of 0.2 V and 0.4 V. Use $\mu_n = 450\text{ cm}^2/\text{V}\cdot\text{s}$.

10.17 It is required to calculate the intrinsic gain A_0 and the unity-gain frequency f_T of an n-channel transistor fabricated in a 0.13-μm CMOS process for which $L_{ov} = 0.1\ L$, $\mu_n = 400\text{ cm}^2/\text{V}\cdot\text{s}$, and $V_A' = 5\text{ V}/\mu\text{m}$. The device is operated at $V_{OV} = 0.2$ V. Find A_0 and f_T for devices with $L = L_{\min}$, $2L_{\min}$, $3L_{\min}$, $4L_{\min}$, and $5L_{\min}$. Present your results in a table. (*Hint:* For f_T, use the approximate expression $f_T \simeq \frac{3\mu_n V_{OV}}{4\pi L^2}$.)

10.18 A particular BJT operating at $I_C = 0.5$ mA has $C_\mu = 1$ pF, $C_\pi = 8$ pF, and $\beta = 100$. What are f_T and f_β for this situation?

10.19 For the transistor described in Problem 10.18, C_π includes a relatively constant depletion-layer capacitance

Transistor	I_E(mA)	r_e(Ω)	g_m(mA/V)	r_π (kΩ)	β_0	f_T(MHz)	C_μ(pF)	C_π(pF)	f_β(MHz)
(a)	2				100	500	2		
(b)		25					2	10.7	4
(c)				2.5	100	500		10.7	
(d)	10				100	500	2		
(e)	0.1				100	150	2		
(f)	1				10	500	2		
(g)						800	1	9	80

of 2 pF. If the device is operated at $I_C = 0.25$ mA, what does its f_T become?

10.20 An *npn* transistor is operated at $I_C = 1$ mA and $V_{CB} = 2$ V. It has $\beta_0 = 100$, $V_A = 50$ V, $\tau_F = 30$ ps, $C_{je0} = 20$ fF, $C_{\mu 0} = 30$ fF, $V_{0c} = 0.75$ V, $m_{CBJ} = 0.5$, and $r_x = 100\ \Omega$. Sketch the complete hybrid-π model, and specify the values of all its components. Also, find f_T.

10.21 Measurement of h_{fe} of an *npn* transistor at 50 MHz shows that $|h_{fe}| = 10$ at $I_C = 0.2$ mA and 12 at $I_C = 1.0$ mA. Furthermore, C_μ was measured and found to be 0.1 pF. Find f_T at each of the two collector currents used. What must τ_F and C_{je} be?

10.22 A particular small-geometry BJT has f_T of 10 GHz and $C_\mu = 0.1$ pF when operated at $I_C = 1.0$ mA. What is C_π in this situation? Also, find g_m. For $\beta = 120$, find r_π and f_β.

10.23 For a BJT whose unity-gain bandwidth is 2 GHz and $\beta_0 = 200$, at what frequency does the magnitude of h_{fe} become 40? What is f_β?

***10.24** For a sufficiently high frequency, measurement of the complex input impedance of a BJT having (ac) grounded emitter and collector yields a real part approximating r_x. For what frequency, defined in terms of ω_β, is such an estimate of r_x good to within 10% under the condition that $r_x \le r_\pi/10$?

***10.25** Complete the table entries on the previous page for transistors (a) through (g), under the conditions indicated. Neglect r_x.

Section 10.3: High-Frequency Response of the CS and CE Amplifiers

10.26 In a particular common-source amplifier for which the midband voltage gain between gate and drain (i.e., $-g_m R_L'$) is -39 V/V, the NMOS transistor has $C_{gs} = 1.0$ pF and $C_{gd} = 0.1$ pF. What input capacitance would you expect? For what range of signal-source resistances can you expect the 3-dB frequency to exceed 1 MHz? Neglect the effect of R_G.

D 10.27 In the circuit of Fig. P10.27, the voltage amplifier is ideal (i.e., it has an infinite input resistance and a zero output resistance).

(a) Use the Miller approach to find an expression for the input capacitance C_{in} in terms of A and C.

(b) Use the expression for C_{in} to obtain the transfer function $V_o(s)/V_{\text{sig}}(s)$.

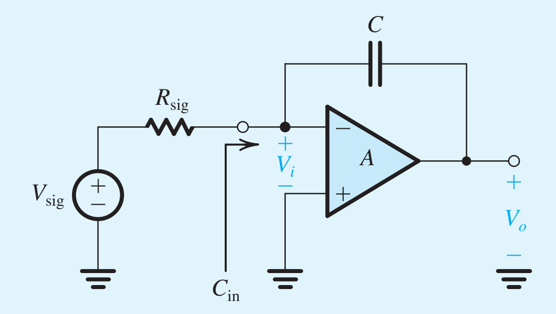


Figure P10.27

(c) If $R_{\text{sig}} = 1\ \text{k}\Omega$, and the gain V_o/V_{sig} is to have a dc value of 40 dB and a 3-dB frequency of 100 kHz, find the values required for A and C.

(d) Sketch a Bode plot for the gain and use it to determine the frequency at which its magnitude reduces to unity.

10.28 An ideal voltage amplifier having a voltage gain of –1000 V/V has a 0.2-pF capacitance connected between its output and input terminals. What is the input capacitance of the amplifier? If the amplifier is fed from a voltage source V_{sig} having a resistance $R_{\text{sig}} = 1\ \text{k}\Omega$, find the transfer function V_o/V_{sig} as a function of the complex-frequency variable s and hence the 3-dB frequency f_H and the unity-gain frequency f_t.

D 10.29 A design is required for a CS amplifier for which the MOSFET is operated at $g_m = 5$ mA/V and has $C_{gs} = 5$ pF and $C_{gd} = 1$ pF. The amplifier is fed with a signal source having $R_{\text{sig}} = 1\ \text{k}\Omega$, and R_G is very large. What is the largest value of R_L' for which the upper 3-dB frequency is at least 6 MHz? What is the corresponding value of midband gain and gain–bandwidth product? If the specification on the upper 3-dB frequency can be relaxed by a factor of 3, that is, to 2 MHz, what can A_M and GB become?

10.30 Reconsider Example 10.3 for the situation in which the transistor is replaced by one whose width W is half that of the original transistor while the bias current remains unchanged. Find modified values for all the device parameters along with A_M, f_H, and the gain–bandwidth product, GB. Contrast this with the original design by calculating the ratios of new value to old for W, V_{OV}, g_m, C_{gs}, C_{gd}, C_{in}, A_M, f_H, and GB.

D *10.31 In a CS amplifier, such as that in Fig. 10.3(a), the resistance of the source $R_{\text{sig}} = 100\ \text{k}\Omega$, amplifier

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

input resistance (which is due to the biasing network) $R_{in} = 100\ \text{k}\Omega$, $C_{gs} = 1\ \text{pF}$, $C_{gd} = 0.2\ \text{pF}$, $g_m = 3\ \text{mA/V}$, $r_o = 50\ \text{k}\Omega$, $R_D = 8\ \text{k}\Omega$, and $R_L = 10\ \text{k}\Omega$. Determine the expected 3-dB cutoff frequency f_H and the midband gain. In evaluating ways to double f_H, a designer considers the alternatives of changing either R_L or R_{in}. To raise f_H as described, what separate change in each would be required? What midband voltage gain results in each case?

10.32 A discrete MOSFET common-source amplifier has $R_G = 2\ \text{M}\Omega$, $g_m = 5\ \text{mA/V}$, $r_o = 100\ \text{k}\Omega$, $R_D = 20\ \text{k}\Omega$, $C_{gs} = 3\ \text{pF}$, and $C_{gd} = 0.5\ \text{pF}$. The amplifier is fed from a voltage source with an internal resistance of 500 kΩ and is connected to a 20-kΩ load. Find:

(a) the overall midband gain A_M
(b) the upper 3-dB frequency f_H
(c) the frequency of the transmission zero, f_Z.

10.33 For the discrete-circuit CS amplifier in Fig. 10.3(a) let $R_{sig} = 100\ \text{k}\Omega$, $R_{G1} = 47\ \text{M}\Omega$, $R_{G2} = 10\ \text{M}\Omega$, $R_S = 2\ \text{k}\Omega$, $R_D = 4.7\ \text{k}\Omega$, $R_L = 10\ \text{k}\Omega$, $g_m = 3\ \text{mA/V}$, $r_o = 100\ \text{k}\Omega$, $C_{gs} = 1\ \text{pF}$, and $C_{gd} = 0.2\ \text{pF}$. Find A_M and f_H.

10.34 Consider the integrated-circuit CS amplifier in Fig. P10.34 for the case $I_{BIAS} = 100\ \mu\text{A}$, Q_2 and Q_3 are matched, and $R_{sig} = 200\ \text{k}\Omega$. For Q_1: $\mu_n C_{ox} = 90\ \mu\text{A/V}^2$, $V_A = 12.8\ \text{V}$, $W/L = 100\ \mu\text{m}/1.6\ \mu\text{m}$, $C_{gs} = 0.2\ \text{pF}$, and $C_{gd} = 0.015\ \text{pF}$. For Q_2: $|V_A| = 19.2\ \text{V}$. Neglecting the effect of the capacitance inevitably present at the output node, find the low-frequency gain, the 3-dB frequency f_H, and the frequency of the zero f_Z.

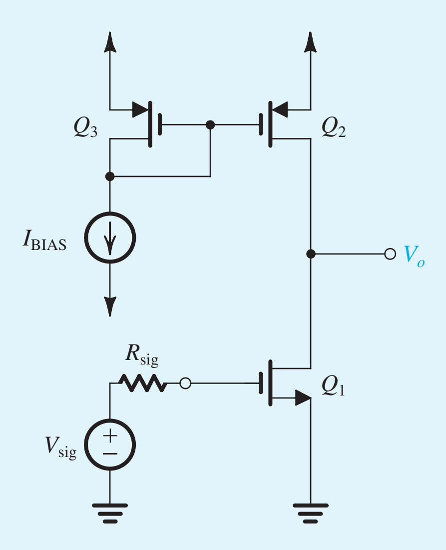


Figure P10.34

10.35 A common-emitter amplifier is measured at midband and found to have a gain of −50 V/V between base and collector. If $C_\pi = 10\ \text{pF}$, $C_\mu = 1\ \text{pF}$, and the effective source resistance $R'_{sig} = 5\ \text{k}\Omega$ [refer to Fig. 10.19(b)], find C_{in} and the 3-dB frequency f_H.

10.36 For a CE amplifier represented by the equivalent circuit in Fig. 10.19(a), let $R_{sig} = 10\ \text{k}\Omega$, $R_B = 100\ \text{k}\Omega$, $r_x = 100\ \Omega$, $C_\pi = 10\ \text{pF}$, $C_\mu = 1\ \text{pF}$, $g_m = 40\ \text{mA/V}$, $r_o = 100\ \text{k}\Omega$, $R_C = 10\ \text{k}\Omega$, $R_L = 10\ \text{k}\Omega$, and $\beta = 100$. Find the midband gain and the 3-dB frequency f_H.

10.37 A designer wishes to investigate the effect of changing the bias current I_E on the midband gain and high-frequency response of the CE amplifier considered in Example 10.4. Let I_E be doubled to 2 mA, and assume that β_0 and f_T remain unchanged at 100 and 800 MHz, respectively. To keep the node voltages nearly unchanged, the designer reduces R_B and R_C by a factor of 2, to 50 kΩ and 4 kΩ, respectively. Assume $r_x = 50\ \Omega$, and recall that $V_A = 100\ \text{V}$ and that C_μ remains constant at 1 pF. As before, the amplifier is fed with a source having $R_{sig} = 5\ \text{k}\Omega$ and feeds a load $R_L = 5\ \text{k}\Omega$. Find the new values of A_M, f_H, and the gain–bandwidth product, $|A_M| f_H$. Comment on the results. Note that the price paid for whatever improvement in performance is achieved is an increase in power. By what factor does the power dissipation increase?

***10.38** The purpose of this problem is to investigate the high-frequency response of the CE amplifier when it is fed with a relatively large source resistance R_{sig}. Refer to the amplifier in Fig. 10.9(a) and to its high-frequency, equivalent-circuit model and the analysis shown in Fig. 10.19. Let $R_B \gg R_{sig}$, $r_x \ll R_{sig}$, $R_{sig} \gg r_\pi$, $g_m R'_L \gg 1$, and $g_m R'_L C_\mu \gg C_\pi$. Under these conditions, show that:

(a) the midband gain $A_M \simeq -\beta R'_L / R_{sig}$
(b) the upper 3-dB frequency $f_H \simeq 1/2\pi C_\mu \beta R'_L$
(c) the gain–bandwidth product $|A_M| f_H \simeq 1/2\pi C_\mu R_{sig}$

Evaluate this approximate value of the gain–bandwidth product for the case $R_{sig} = 25\ \text{k}\Omega$ and $C_\mu = 1\ \text{pF}$. Now, if the transistor is biased at $I_C = 1\ \text{mA}$ and has $\beta = 100$, find the midband gain and f_H for the two cases $R'_L = 25\ \text{k}\Omega$ and $R'_L = 2.5\ \text{k}\Omega$. On the same coordinates, sketch Bode plots for the gain magnitude versus frequency for the two cases. What f_H is obtained when the gain is unity? What value of R'_L corresponds?

10.39 For a version of the CE amplifier circuit in Fig. 10.9(a), $R_{sig} = 10\ \text{k}\Omega$, $R_{B1} = 68\ \text{k}\Omega$, $R_{B2} = 27\ \text{k}\Omega$,

$R_E = 2.2\ \text{k}\Omega$, $R_C = 4.7\ \text{k}\Omega$, and $R_L = 10\ \text{k}\Omega$. The collector current is 0.8 mA, $\beta = 200$, $f_T = 1$ GHz, and $C_\mu = 0.8$ pF. Neglecting the effect of r_x and r_o, find the midband voltage gain and the upper 3-dB frequency f_H.

10.40 Consider an ideal voltage amplifier with a gain of 0.9 V/V, and a resistance $R = 100\ \text{k}\Omega$ connected in the feedback path—that is, between the output and input terminals. Use Miller's theorem to find the input resistance of this circuit.

10.41 The amplifiers listed below are characterized by the descriptor (A, C), where A is the voltage gain from input to output and C is an internal capacitor connected between input and output. For each, find the equivalent capacitances at the input and at the output as provided by the use of Miller's theorem:

(a) –1000 V/V, 1 pF
(b) –10 V/V, 10 pF
(c) –1 V/V, 10 pF
(d) +1 V/V, 10 pF
(e) +10 V/V, 10 pF

Note that the input capacitance found in case (e) can be used to cancel the effect of other capacitance connected from input to ground. In (e), what capacitance can be canceled?

***10.42** Figure P10.42 shows an ideal voltage amplifier with a gain of +2 V/V (usually implemented with an op amp connected in the noninverting configuration) and a resistance R connected between output and input.

(a) Using Miller's theorem, show that the input resistance $R_{in} = -R$.
(b) Use Norton's theorem to replace V_{sig}, R_{sig}, and R_{in} with a signal current source and an equivalent parallel resistance. Show that by selecting $R_{sig} = R$, the equivalent parallel resistance becomes infinite and the current I_L into the load impedance Z_L becomes V_{sig}/R. The circuit then functions as an ideal voltage-controlled current source with an output current I_L.
(c) If Z_L is a capacitor C, find the transfer function V_o/V_{sig} and show it is that of an ideal noninverting integrator.

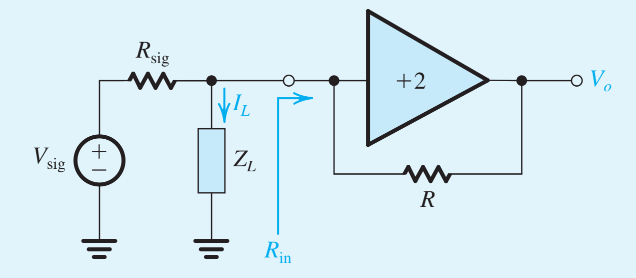


Figure P10.42

10.43 Use Miller's theorem to investigate the performance of the inverting op-amp circuit shown in Fig. P10.43. Assume the op amp to be ideal except for having a finite differential gain, A. Without using any knowledge of op-amp circuit analysis, find R_{in}, V_i, V_o, and V_o/V_{sig}, for each of the following values of A: 10 V/V, 100 V/V, 1000 V/V, and 10,000 V/V. Assume $V_{sig} = 1$ V. Present your results in the table below.

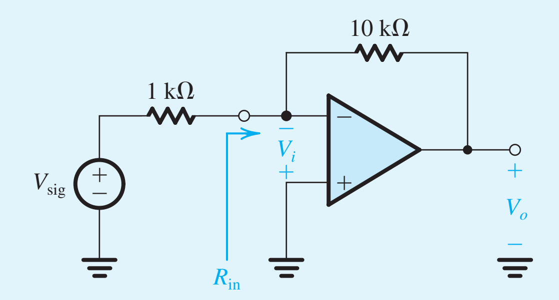


Figure P10.43

A	R_{in}	V_i	V_o	V_o/V_{sig}
10 V/V				
100 V/V				
1000 V/V				
10,000 V/V				

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

***10.44** The amplifier shown in Fig. P10.44 has $R_{sig} = R_L = 1\ k\Omega$, $R_C = 1\ k\Omega$, $R_B = 47\ k\Omega$, $\beta = 100$, $C_\mu = 0.8$ pF, and $f_T = 600$ MHz. Assume the coupling capacitors to be very large.

(a) Find the dc collector current of the transistor.
(b) Find g_m and r_π.
(c) Neglecting r_o, find the midband voltage gain from base to collector (neglect the effect of R_B).
(d) Use the gain obtained in (c) to find the component of R_{in} that arises as a result of R_B. Hence find R_{in}.
(e) Find the overall gain at midband.
(f) Find C_{in}.
(g) Find f_H.

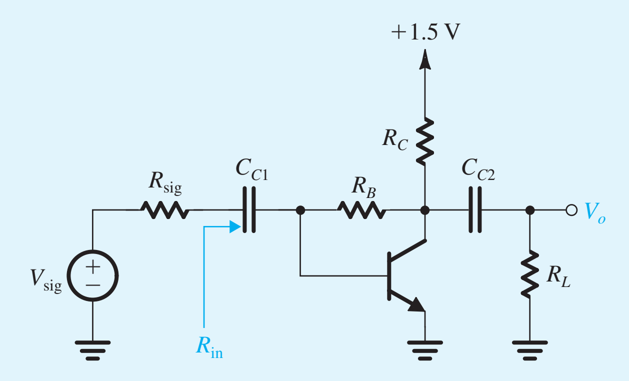


Figure P10.44

***10.45** Figure P10.45 shows a diode-connected transistor with the bias circuit omitted. Utilizing the BJT high-frequency, hybrid-π model with $r_x = 0$ and $r_o = \infty$, derive an expression for $Z_i(s)$ as a function of r_e and C_π. Find the frequency at which the impedance has a phase angle of 45° for the case in which the BJT has $f_T = 400$ MHz and the bias current is relatively high. What is the frequency when the bias current is reduced so that $C_\pi \simeq C_\mu$? Assume $\alpha = 1$.

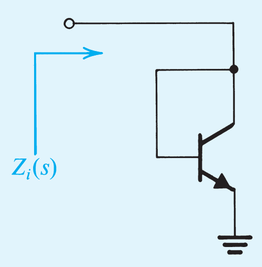


Figure P10.45

10.46 A CS amplifier modeled with the equivalent circuit of Fig. 10.22(a) is specified to have $C_{gs} = 2$ pF, $C_{gd} = 0.1$ pF, $g_m = 4$ mA/V, $C_L = 2$ pF, and $R'_L = 20\ k\Omega$. Find A_M, f_{3dB}, f_Z, and f_t.

D 10.47 A common-source amplifier fed with a low-resistance signal source and operating with $g_m = 2$ mA/V has a unity-gain frequency of 2 GHz. What additional capacitance must be connected to the drain node to reduce f_t to 1 GHz?

***10.48** It is required to analyze the high-frequency response of the CMOS amplifier shown in Fig. P10.34 for the case $R_{sig} = 0$. The dc bias current is 100 μA. For Q_1, $\mu_n C_{ox} = 90\ \mu A/V^2$, $V_A = 12.8$ V, $W/L = 100\ \mu m/1.6\ \mu m$, $C_{gs} = 0.2$ pF, $C_{gd} = 0.015$ pF, and $C_{db} = 20$ fF. For Q_2, $C_{gd} = 0.015$ pF, $C_{db} = 36$ fF, and $|V_A| = 19.2$ V. For simplicity, assume that the signal voltage at the gate of Q_2 is zero. Find the low-frequency gain, the frequency of the pole, and the frequency of the zero. (*Hint:* The total capacitance at the output mode $= C_{db1} + C_{db2} + C_{gd2}$).

10.49 Consider an active-loaded common-emitter amplifier. Let the amplifier be fed with an ideal voltage source V_i, and neglect the effect of r_x. Assume that the load current source has a very high resistance and that there is a capacitance C_L present between the output node and ground. This capacitance represents the sum of the input capacitance of the subsequent stage and the inevitable parasitic capacitance between collector and ground. Show that the voltage gain is given by

$$\frac{V_o}{V_i} = -g_m r_o \frac{1 - s(C_\mu / g_m)}{1 + s(C_L + C_\mu) r_o}$$

If the transistor is biased at $I_C = 200\ \mu A$ and $V_A = 100$ V, $C_\mu = 0.2$ pF, and $C_L = 1$ pF, find the dc gain, the 3-dB frequency, the frequency of the zero, and the frequency at which the gain reduces to unity. Sketch a Bode plot for the gain magnitude.

10.50 A particular BJT operating at 2 mA is specified to have $f_T = 2$ GHz, $C_\mu = 1$ pF, $r_x = 100\ \Omega$, and $\beta = 120$. The device is used in a CE amplifier operating from a very-low-resistance voltage source.

(a) If the midband gain obtained is −10 V/V, what is the value of f_H?
(b) If the midband gain is reduced to −1 V/V (by changing R'_L), what f_H is obtained?

Section 10.4: Useful Tools for the Analysis of the High-Frequency Response of Amplifiers

10.51 A direct-coupled amplifier has a low-frequency gain of 40 dB, poles at 2 MHz and 20 MHz, a zero on the negative real axis at 200 MHz, and another zero at infinite frequency. Express the amplifier gain function in the form of Eqs. (10.70) and (10.71), and sketch a Bode plot for the gain magnitude. What do you estimate the 3-dB frequency f_H to be?

10.52 An amplifier with a dc gain of 60 dB has a single-pole, high-frequency response with a 3-dB frequency of 100 kHz.

(a) Give an expression for the gain function $A(s)$.
(b) Sketch Bode diagrams for the gain magnitude and phase.
(c) What is the gain–bandwidth product?
(d) What is the unity-gain frequency?
(e) If a change in the amplifier circuit causes its transfer function to acquire another pole at 1 MHz, sketch the resulting gain magnitude and specify the unity-gain frequency. Note that this is an example of an amplifier with a unity-gain bandwidth that is different from its gain–bandwidth product.

10.53 Consider an amplifier whose $F_H(s)$ is given by

$$F_H(s) = \frac{1}{\left(1+\dfrac{s}{\omega_{P1}}\right)\left(1+\dfrac{s}{\omega_{P2}}\right)}$$

with $\omega_{P1} < \omega_{P2}$. Find the ratio ω_{P2}/ω_{P1} for which the value of the 3-dB frequency ω_H calculated using the dominant-pole approximation differs from that calculated using the root-sum-of-squares formula (Eq. 10.77) by:

(a) 10%
(b) 1%

10.54 The high-frequency response of a direct-coupled amplifier having a dc gain of –1000 V/V incorporates zeros at ∞ and 10^4 rad/s (one at each frequency) and poles at 10^3 rad/s and 10^5 rad/s (one at each frequency). Write an expression for the amplifier transfer function. Find ω_H using

(a) the dominant-pole approximation
(b) the root-sum-of-squares approximation (Eq. 10.77).

If a way is found to lower the frequency of the finite zero to 10^3 rad/s, what does the transfer function become? What is the 3-dB frequency of the resulting amplifier?

10.55 A direct-coupled amplifier has a dominant pole at 1000 rad/s and three coincident poles at a much higher frequency. These nondominant poles cause the phase lag of the amplifier at high frequencies to exceed the 90° angle due to the dominant pole. It is required to limit the excess phase at $\omega = 10^7$ rad/s to 30° (i.e., to limit the total phase angle to –120°). Find the corresponding frequency of the nondominant poles.

10.56 An IC CS amplifier has $g_m = 2$ mA/V, $C_{gs} = 30$ fF, $C_{gd} = 5$ fF, $C_L = 30$ fF, $R'_{sig} = 10$ kΩ, and $R'_L = 20$ kΩ. Use the method of open-circuit time constants to obtain an estimate for f_H. Also, find the frequency of the transmission zero, f_Z.

10.57 For a particular amplifier modeled by the circuit of Fig. 10.18(a), $g_m = 5$ mA/V, $R_{sig} = 150$ kΩ, $R_G = 0.65$ MΩ, $R'_L = 10$ kΩ, $C_{gs} = 2$ pF, and $C_{gd} = 0.5$ pF. There is also a load capacitance of 30 pF. Find the corresponding midband voltage gain, the open-circuit time constants, and an estimate of the 3-dB frequency.

10.58 Consider the high-frequency response of an amplifier consisting of two identical stages in cascade, each with an input resistance of 10 kΩ and an output resistance of 2 kΩ. The two-stage amplifier is driven from a 10-kΩ source and drives a 1-kΩ load. Associated with each stage is a parasitic input capacitance (to ground) of 10 pF and a parasitic output capacitance (to ground) of 2 pF. Parasitic capacitances of 10 pF and 7 pF also are associated with the signal-source and load connections, respectively. For this arrangement, find the three poles and estimate the 3-dB frequency f_H.

10.59 A CS amplifier that can be represented by the equivalent circuit of Fig. 10.24 has $C_{gs} = 2$ pF, $C_{gd} = 0.1$ pF, $C_L = 2$ pF, $g_m = 4$ mA/V, and $R'_{sig} = R'_L = 20$ kΩ. Find the midband gain A_M, the input capacitance C_{in} using the Miller approximation, and hence an estimate of the 3-dB frequency f_H. Also, obtain another estimate of f_H using open-circuit time constants. Which of the two estimates is more appropriate and why?

D 10.60 For a CS amplifier with $g_m = 5$ mA/V, $C_{gs} = 5$ pF, $C_{gd} = 1$ pF, $C_L = 5$ pF, $R'_{sig} = 10$ kΩ, and $R'_L = 10$ kΩ, find τ_H and f_H. What is the percentage of τ_H that is caused by the interaction of R'_{sig} with the input capacitance? To what value must R'_{sig} be lowered in order to double f_H?

D 10.61 For the CS amplifier in Example 10.8, find the value of the additional capacitance to be connected at the output node in order to lower f_H to 100 MHz.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

10.62 Consider the CE amplifier whose equivalent circuit is shown in Fig. 10.19(a) but with a capacitance C_L connected across the output terminals. Let $R'_{sig} = 5\text{ k}\Omega$, $R_B = \infty$, $r_x = 0$, $g_m = 20\text{ mA/V}$, $\beta = 100$, $C_\pi = 10\text{ pF}$, $C_\mu = 1\text{ pF}$, $R'_L = 5\text{ k}\Omega$, and $C_L = 10\text{ pF}$. Find A_M and f_H.

10.63 A common-emitter amplifier has $C_\pi = 10\text{ pF}$, $C_\mu = 0.3\text{ pF}$, $C_L = 3\text{ pF}$, $g_m = 40\text{ mA/V}$, $\beta = 100$, $r_x = 100\ \Omega$, $R'_L = 5\text{ k}\Omega$, and $R_{sig} = 1\text{ k}\Omega$. Find the midband gain A_M and an estimate of the 3-dB frequency f_H using the Miller approximation. Also, obtain another estimate of f_H using the method of open-circuit time constants. Which of the two estimates would you consider to be more realistic, and why?

10.64 Consider a CS amplifier loaded in a current source with an output resistance equal to r_o of the amplifying transistor. The amplifier is fed from a signal source with $R_{sig} = r_o/2$. The transistor is biased to operate at $g_m = 2\text{ mA/V}$ and $r_o = 20\text{ k}\Omega$; $C_{gs} = C_{gd} = 0.1\text{ pF}$. Use the Miller approximation to determine an estimate of f_H. Repeat for the following two cases: (i) the bias current I in the entire system is reduced by a factor of 4, and (ii) the bias current I in the entire system is increased by a factor of 4. Remember that both R_{sig} and R_L will change as r_o changes.

10.65 Use the method of open-circuit time constants to find f_H for a CS amplifier for which $g_m = 1.5\text{ mA/V}$, $C_{gs} = C_{gd} = 0.2\text{ pF}$, $r_o = 20\text{ k}\Omega$, $R_L = 12\text{ k}\Omega$, and $R_{sig} = 100\text{ k}\Omega$ for the following cases: (a) $C_L = 0$, (b) $C_L = 10\text{ pF}$, and (c) $C_L = 50\text{ pF}$. Compare with the value of f_H obtained using the Miller approximation.

Section 10.5: High-Frequency Response of the Common-Gate and Cascode Amplifiers

10.66 A CG amplifier is specified to have $C_{gs} = 4\text{ pF}$, $C_{gd} = 0.2\text{ pF}$, $C_L = 2\text{ pF}$, $g_m = 5\text{ mA/V}$, $R_{sig} = 1\text{ k}\Omega$, and $R_L = 10\text{ k}\Omega$. Neglecting the effects of r_o, find the low-frequency gain V_o/V_{sig}, the frequencies of the poles f_{P1} and f_{P2}, and hence an estimate of the 3-dB frequency f_H.

***10.67** Sketch the high-frequency equivalent circuit of a CB amplifier fed from a signal generator characterized by V_{sig} and R_{sig} and feeding a load resistance R_L in parallel with a capacitance C_L.

(a) Show that for $r_x = 0$ and $r_o = \infty$, the circuit can be separated into two parts: an input part that produces a pole at

$$f_{P1} = \frac{1}{2\pi C_\pi (R_{sig} \| r_e)}$$

and an output part that forms a pole at

$$f_{P2} = \frac{1}{2\pi (C_\mu + C_L) R_L}$$

Note that these are the bipolar counterparts of the MOS expressions in Eqs. (10.94) and (10.95).

(b) Evaluate f_{P1} and f_{P2} and hence obtain an estimate for f_H for the case $C_\pi = 10\text{ pF}$, $C_\mu = 1\text{ pF}$, $C_L = 1\text{ pF}$, $I_C = 1\text{ mA}$, $R_{sig} = 1\text{ k}\Omega$, and $R_L = 10\text{ k}\Omega$. Also, find f_T of the transistor.

10.68 Consider a CG amplifier loaded in a resistance $R_L = r_o$ and fed with a signal source having a resistance $R_{sig} = r_o/2$. Also let $C_L = C_{gs}$. Use the method of open-circuit time constants to show that for $g_m r_o \gg 1$, the upper 3-dB frequency is related to the MOSFET f_T by the approximate expression

$$f_H = f_T/(g_m r_o)$$

10.69 For the CG amplifier in Example 10.9, how much additional capacitance should be connected between the output node and ground to reduce f_H to 200 MHz?

10.70 An IC CG amplifier is fed from a signal source with $R_{sig} = r_o/2$, where r_o is the MOSFET output resistance. It has a current-source load with an output resistance equal to r_o. The MOSFET is operated at $I_D = 100\ \mu\text{A}$ and has $g_m = 1.5\text{ mA/V}$, $V_A = 10\text{ V}$, $C_{gs} = 0.2\text{ pF}$, $C_{gd} = 0.015\text{ pF}$, and $C_{db} = 20\text{ fF}$. As well, the current-source load provides an additional 30 fF capacitance at the output node. Find f_H.

10.71 Find the dc gain and the 3-dB frequency of a MOS cascode amplifier operated at $g_m = 2\text{ mA/V}$ and $r_o = 20\text{ k}\Omega$. The MOSFETs have $C_{gs} = 20\text{ fF}$, $C_{gd} = 5\text{ fF}$, and $C_{db} = 5\text{ fF}$. The amplifier is fed from a signal source with $R_{sig} = 100\text{ k}\Omega$ and is connected to a load resistance of 1 MΩ. There is also a load capacitance C_L of 20 fF.

***10.72** (a) Consider a CS amplifier having $C_{gd} = 0.2\text{ pF}$, $R_{sig} = R_L = 20\text{ k}\Omega$, $g_m = 4\text{mA/V}$, $C_{gs} = 2\text{ pF}$, C_L (including C_{db}) $= 1\text{ pF}$, $C_{db} = 0.2\text{ pF}$, and $r_o = 20\text{ k}\Omega$. Find the low-frequency gain A_M, and estimate f_H using open-circuit time constants. Hence determine the gain–bandwidth product. (b) If a CG stage utilizing an identical MOSFET is cascaded with the CS transistor in (a) to create a cascode amplifier, determine the new values of A_M, f_H, and gain–bandwidth product. Assume R_L remains unchanged.

D 10.73 It is required to design a cascode amplifier to provide a dc gain of 74 dB when driven with a low-resistance generator and utilizing NMOS transistors for which $V_A = 10$ V, $\mu_n C_{ox} = 200\ \mu A/V^2$, $W/L = 50$, $C_{gd} = 0.1$ pF, and $C_L = 1$ pF. Assuming that $R_L = R_o$, determine the overdrive voltage and the drain current at which the MOSFETs should be operated. Find the unity-gain frequency and the 3-dB frequency. If the cascode transistor is removed and R_L remains unchanged, what will the dc gain become?

10.74 (a) Show that introducing a cascode transistor to an IC CS amplifier whose bandwidth is limited by the interaction of R_{sig} and the input capacitance, and whose load resistance is equal to r_o, increases the dc gain by approximately a factor of 2 and f_H by the factor N,

$$N = \frac{C_{gs} + \frac{1}{2}(g_m r_o)C_{gd}}{C_{gs} + 3C_{gd}}$$

Assume that the bandwidth of the cascode amplifier is primarily determined by the input circuit.

(b) If $C_{gd} = 0.1\,C_{gs}$ and the dc gain of the CS amplifier is 50, what is the value of N?

(c) If $V_A = 10$ V, $\mu_n C_{ox} = 400\ \mu A/V^2$, and $W/L = 10$, find V_{OV} and I_D at which the transistors must be operating.

10.75 (a) For an integrated-circuit MOS cascode amplifier fed with a source having a very small resistance and loaded in a resistance equal to its R_o, use the expression for the unity-gain bandwidth in Fig. 10.29 to show that

$$f_t = \frac{\sqrt{2\mu_n C_{ox}(W/L)}}{2\pi(C_L + C_{gd})}\sqrt{I_D}$$

(b) For $\mu_n C_{ox} = 400\ \mu A/V^2$, $W/L = 20$, $C_L = 20$ fF, $C_{gd} = 5$ fF, and $V_A = 10$ V, provide in table form f_t (GHz), V_{OV} (V), g_m (mA/V), r_o (kΩ), R_o (MΩ), A_M (V/V), and f_H (MHz) for $I_D = 100\ \mu A$, 200 μA, and 500 μA.

10.76 Consider a bipolar cascode amplifier biased at a current of 1 mA. The transistors used have $\beta = 100$, $r_o = 100$ kΩ, $C_\pi = 10$ pF, $C_\mu = 2$ pF, $C_{cs} = 0$, and $r_x = 50\ \Omega$. The amplifier is fed with a signal source having $R_{sig} = 5$ kΩ. The load resistance $R_L = 2$ kΩ. Find the low-frequency gain A_M, and estimate the value of the 3-dB frequency f_H.

***10.77** In this problem we consider the frequency response of the bipolar cascode amplifier in the case that r_o can be neglected.

(a) Refer to the circuit in Fig. 10.30, and note that the total resistance between the collector of Q_1 and ground will be equal to r_{e2}, which is usually very small. It follows that the pole introduced at this node will typically be at a very high frequency and thus will have negligible effect on f_H. It also follows that at the frequencies of interest the gain from the base to the collector of Q_1 will be $-g_{m1}r_{e2} \simeq -1$. Use this to find the capacitance at the input of Q_1 and hence show that the pole introduced at the input node will have a frequency

$$f_{P1} \simeq \frac{1}{2\pi R'_{sig}(C_{\pi 1} + 2C_{\mu 1})}$$

Then show that the pole introduced at the output node will have a frequency

$$f_{P2} \simeq \frac{1}{2\pi R_L(C_L + C_{cs2} + C_{\mu 2})}$$

(b) Evaluate f_{P1} and f_{P2}, and use the sum-of-the-squares formula to estimate f_H for the amplifier with $I = 1$ mA, $C_\pi = 10$ pF, $C_\mu = 2$ pF, $C_{cs} = C_L = 0$, $\beta = 100$, $R_L = 2$ kΩ, and $r_x = 0$ in the following two cases:
(i) $R_{sig} = 1$ kΩ
(ii) $R_{sig} = 10$ kΩ

10.78 A BJT cascode amplifier uses transistors for which $\beta = 100$, $V_A = 100$ V, $f_T = 1$ GHz, and $C_\mu = 0.1$ pF. It operates at a bias current of 0.1 mA between a source with $R_{sig} = r_\pi$ and a load $R_L = \beta r_o$. Let $C_L = C_{cs} = 0$, and $r_x = 0$. Find the overall voltage gain at dc. By evaluating the various components of τ_H show that the pole introduced at the output mode is dominant. Find its frequency and hence an estimate of f_H and f_t.

Section 10.6: High-Frequency Response of the Source and Emitter Followers

10.79 A source follower has $g_m = 5$ mA/V, $g_{mb} = 0$, $r_o = 20$ kΩ, $R_{sig} = 20$ kΩ, $R_L = 2$ kΩ, $C_{gs} = 2$ pF, $C_{gd} = 0.1$ pF, and $C_L = 1$ pF. Find A_M, R_o, f_Z, the frequencies of the two poles, and an estimate of f_H.

10.80 Using the expression for the source follower f_H in Eq. (10.124) show that for situations in which $C_L = 0$, R_{sig}

is large and R_L is small,

$$f_H \simeq \frac{1}{2\pi R_{sig}\left(C_{gd} + \dfrac{C_{gs}}{1+g_m R_L'}\right)}$$

Find f_H for the case $R_{sig} = 100\text{ k}\Omega$, $R_L = 2\text{ k}\Omega$, $r_o = 20\text{ k}\Omega$, $g_m = 5\text{ mA/V}$, $C_{gd} = 10\text{ pF}$, and $C_{gs} = 2\text{ pF}$.

10.81 Refer to Fig. 10.31(c). In situations in which R_{sig} is large, the high-frequency response of the source follower is determined by the low-pass circuit formed by R_{sig} and the input capacitance. An estimate of C_{in} can be obtained by using the Miller approximation to replace C_{gs} with an input capacitance $C_{eq} = C_{gs}(1-K)$ where K is the gain from gate to source. Using the low-frequency value of $K = g_m R_L'/(1+g_m R_L')$ find C_{eq} and hence C_{in} and an estimate of f_H.

10.82 A source follower has a maximally flat gain response with a dc gain of 0.8 and a 3-dB frequency of 1 MHz. Give its transfer function.

10.83 A discrete-circuit source follower driven with $R_{sig} = 100\text{ k}\Omega$ has $C_{gs} = 10\text{ pF}$, $C_{gd} = 1\text{ pF}$, $C_L = 10\text{ pF}$, $g_{mb} = 0$, and r_o very large. The transfer function of the source follower is measured as R_L is varied. At what value of R_L will the transfer function be maximally flat? At this value of R_L the dc gain is found to be 0.9 V/V. What is the 3-dB frequency? What is the value of g_m at which the source follower is operating?

10.84 For an emitter follower biased at $I_C = 1\text{ mA}$, having $R_{sig} = R_L = 1\text{ k}\Omega$, and using a transistor specified to have $f_T = 2\text{ GHz}$, $C_\mu = 0.1\text{ pF}$, $C_L = 0$, $r_x = 100\ \Omega$, $\beta = 100$, and $V_A = 20\text{ V}$, evaluate the low-frequency gain A_M, the frequency of the transmission zero, the pole frequencies, and an estimate of the 3-dB frequency f_H.

Section 10.7: High-Frequency Response of Differential Amplifiers

10.85 A MOSFET differential amplifier such as that shown in Fig. 10.34(a) is biased with a current source $I = 400\ \mu\text{A}$. The transistors have $W/L = 16$, $k_n' = 400\ \mu\text{A/V}^2$, $V_A = 20\text{ V}$, $C_{gs} = 40\text{ fF}$, $C_{gd} = 5\text{ fF}$, and $C_{db} = 5\text{ fF}$. The drain resistors are $10\text{ k}\Omega$ each. Also, there is a 100-fF capacitive load between each drain and ground.

(a) Find V_{OV} and g_m for each transistor.
(b) Find the differential gain A_d.
(c) If the input signal source has a small resistance R_{sig} and thus the frequency response is determined primarily by the output pole, estimate the 3-dB frequency f_H.
(d) If, in a different situation, the amplifier is fed symmetrically with a signal source of $40\text{ k}\Omega$ resistance (i.e., $20\text{ k}\Omega$ in series with each gate terminal), use the open-circuit time-constants method to estimate f_H.

10.86 A MOS differential amplifier is biased with a current source having an output resistance $R_{SS} = 100\text{ k}\Omega$ and an output capacitance $C_{SS} = 1\text{ pF}$. If the differential gain is found to have a dominant pole at 20 MHz, what is the 3-dB frequency of the CMRR?

10.87 The differential gain of a MOS amplifier is 100 V/V with a dominant pole at 10 MHz. The common-mode gain is 0.1 V/V at low frequencies and has a transmission zero at 1 MHz. Sketch a Bode plot for the CMRR.

10.88 In a particular MOS differential amplifier design, the bias current $I = 100\ \mu\text{A}$ is provided by a single transistor operating at $V_{OV} = 0.4\text{ V}$ with $V_A = 40\text{ V}$ and output capacitance C_{SS} of 100 fF. What is the frequency of the common-mode gain zero (f_Z) at which A_{cm} begins to rise above its low-frequency value? To meet a requirement for reduced power supply, consideration is given to reducing V_{OV} to 0.2 V while keeping I unchanged. Assuming the current-source capacitance to be directly proportional to the device width, what is the impact on f_Z of this proposed change?

10.89 Repeat Exercise 10.26 for the situation in which the bias current is reduced to $80\ \mu\text{A}$ and R_D is raised to $20\text{ k}\Omega$. For (d), let R_{sig} be raised from $20\text{ k}\Omega$ to $100\text{ k}\Omega$. (*Note*: This is a low-voltage, low-power design.)

10.90 A BJT differential amplifier operating with a 0.5-mA current source uses transistors for which $\beta = 100$, $f_T = 500$ MHz, $C_\mu = 0.5\text{ pF}$, and $r_x = 100\ \Omega$. Each of the collector resistances is $10\text{ k}\Omega$, and r_o is very large. The amplifier is fed in a symmetrical fashion with a source resistance of $10\text{ k}\Omega$ in series with each of the two input terminals.

(a) Sketch the differential half-circuit and its high-frequency equivalent circuit.
(b) Determine the low-frequency value of the overall differential gain.
(c) Use the Miller approximation to determine the input capacitance and hence estimate the 3-dB frequency f_H and the gain–bandwidth product.

10.91 A differential amplifier is biased by a current source having an output resistance of 1 MΩ and an output capacitance of 1 pF. The differential gain exhibits a dominant pole at 2 MHz. What are the poles of the CMRR?

10.92 A current-mirror-loaded MOS differential amplifier is biased with a current source $I = 0.2$ mA. The two NMOS transistors of the differential pair are operating at $V_{OV} = 0.2$ V, and the PMOS devices of the mirror are operating at $|V_{OV}| = 0.2$ V. The Early voltage $V_{An} = |V_{Ap}| = 10$ V. The total capacitance at the input node of the mirror is 0.1 pF and that at the output node of the amplifier is 0.2 pF. Find the dc value and the frequencies of the poles and zero of the differential voltage gain.

10.93 Consider the current-mirror-loaded CMOS differential amplifier of Fig. 10.37(a) for the case of all transistors operated at the same $|V_{OV}|$ and having the same $|V_A|$. Also let the total capacitance at the output node (C_L) be four times the total capacitance at the input node of the current mirror C_m. Give expressions for A_d, f_{P1}, f_{P2}, and f_Z. Hence show that $f_{P2}/f_{P1} = 4A_d$ and $f_t = g_m/2\pi C_L$. For $V_A = 20$ V, $V_{OV} = 0.2$ V, $I = 0.2$ mA, $C_L = 100$ fF, and $C_m = 25$ fF, find the dc value of A_d, and the value of f_{P1}, f_t, f_{P2}, and f_Z and sketch a Bode plot for $|A_d|$.

***10.94** For the current mirror in Fig. P10.94, derive an expression for the current transfer function $I_o(s)/I_i(s)$ taking into account the BJT internal capacitances and neglecting r_x and r_o. Assume the BJTs to be identical. Observe that a signal ground appears at the collector of Q_2. If the mirror is biased at 1 mA and the BJTs at this operating point are characterized by $f_T = 500$ MHz, $C_\mu = 2$ pF, and $\beta_0 = 100$, find the frequencies of the pole and zero of the transfer function.

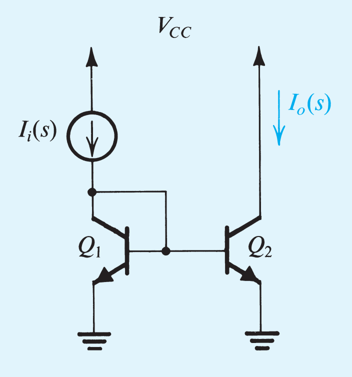


Figure P10.94

Section 10.8: Other Wideband Amplifier Configurations

10.95 Consider the case of a discrete-circuit CS amplifier in which a source-degeneration resistance is utilized to control the bandwidth. Assume that r_o is very large and C_L is negligibly small. Adapt the formulas given in the text for this case and thus give the expressions for A_M and f_H. Let $R_{sig} = 100$ kΩ, $g_m = 5$ mA/V, $R_L = 5$ kΩ, $C_{gs} = 10$ pF, and $C_{gd} = 2$ pF. Find $|A_M|$, f_H, and the gain–bandwidth product for these three cases: $R_s = 0$, 100 Ω, and 200 Ω.

10.96 A CS amplifier is specified to have $g_m = 5$ mA/V, $r_o = 40$ kΩ, $C_{gs} = 2$ pF, $g_{gd} = 0.1$ pF, $C_L = 1$ pF, $R_{sig} = 20$ kΩ, and $R_L = 40$ kΩ.

(a) Find the low-frequency gain A_M, and use open-circuit time constants to estimate the 3-dB frequency f_H. Hence determine the gain–bandwidth product.
(b) If a 400-Ω resistance is connected in the source lead, find the new values of $|A_M|$, f_H, and the gain–bandwidth product.

D 10.97 (a) Use the approximate expression in Eq. (10.156) to determine the gain–bandwidth product of a CS amplifier with a source-degeneration resistance. Assume $C_{gd} = 0.2$ pF and $R_{sig} = 100$ kΩ.
(b) If a low-frequency gain of 20 V/V is required, what f_H corresponds?
(c) For $g_m = 5$ mA/V, $A_0 = 100$ V/V, and $R_L = 20$ kΩ, find the required value of R_s.

10.98 For the CS amplifier with a source-degeneration resistance R_s, show for $R_{sig} \gg R_s$, $r_o \gg R_s$, and $R_L = r_o$ that

$$A_M = \frac{-A_0}{2+k}$$

and

$$\tau_H \simeq \frac{C_{gs}R_{sig}}{1+(k/2)} + C_{gd}R_{sig}\left(1+\frac{A_0}{2+k}\right) + (C_L + C_{gd})r_o\left(\frac{1+k}{2+k}\right)$$

where $k \equiv g_m R_s$

D *10.99 It is required to generate a table of $|A_M|$, f_H, and f_t versus $k \equiv g_m R_s$ for a CS amplifier with a source-degeneration resistance R_s. The table should have entries for $k = 0, 1, 2, \ldots, 15$. The amplifier is specified to have $g_m = 5$ mA/V, $r_o = 40$ kΩ, $R_L = 40$ kΩ, $R_{sig} = 20$ kΩ, $C_{gs} = 2$ pF, $C_{gd} = 0.1$ pF, and $C_L = 1$ pF. Use the formulas

for A_M and τ_H given in the statement for Problem 10.98. If $f_H = 2$ MHz is required, find the value needed for R_s and the corresponding value of $|A_M|$.

***10.100** In this problem we investigate the bandwidth extension obtained by placing a source follower between the signal source and the input of the CS amplifier.

(a) First consider the CS amplifier of Fig. P10.100(a). Show that

$$A_M = -g_m r_o$$

$$\tau_H = C_{gs} R_{\text{sig}} + C_{gd}\left[R_{\text{sig}}(1 + g_m r_o) + r_o\right] + C_L r_o$$

where C_L is the total capacitance between the output node and ground. Calculate the value of A_M, f_H, and the gain–bandwidth product for the case $g_m = 1$ mA/V, $r_o = 20$ kΩ, $R_{\text{sig}} = 20$ kΩ, $C_{gs} = 20$ fF, $C_{gd} = 5$ fF, and $C_L = 10$ fF.

(b) For the CD−CS amplifier in Fig. P10.100(b), show that

$$A_M = -\frac{r_{o1}}{1/g_{m1} + r_{o1}}(g_{m2} r_{o2})$$

$$\tau_H = C_{gd1} R_{\text{sig}} + C_{gs1}\frac{R_{\text{sig}} + r_{o1}}{1 + g_{m1} r_{o1}} + C_{gs2}\left(\frac{1}{g_{m1}} \| r_{o1}\right) + C_{gd2}\left[\left(\frac{1}{g_{m1}} \| r_{o1}\right)(1 + g_{m2} r_{o2}) + r_{o2}\right] + C_L r_{o2}$$

Calculate the values of A_M, f_H, and the gain–bandwidth product for the same parameter values used in (a). Compare with the results of (a).

***10.101** The transistors in the circuit of Fig. P10.101 have $\beta_0 = 100$, $V_A = 100$ V, and $C_\mu = 0.2$ pF. At a bias current of 100 µA, $f_T = 200$ MHz. (Note that the bias details are not shown.)

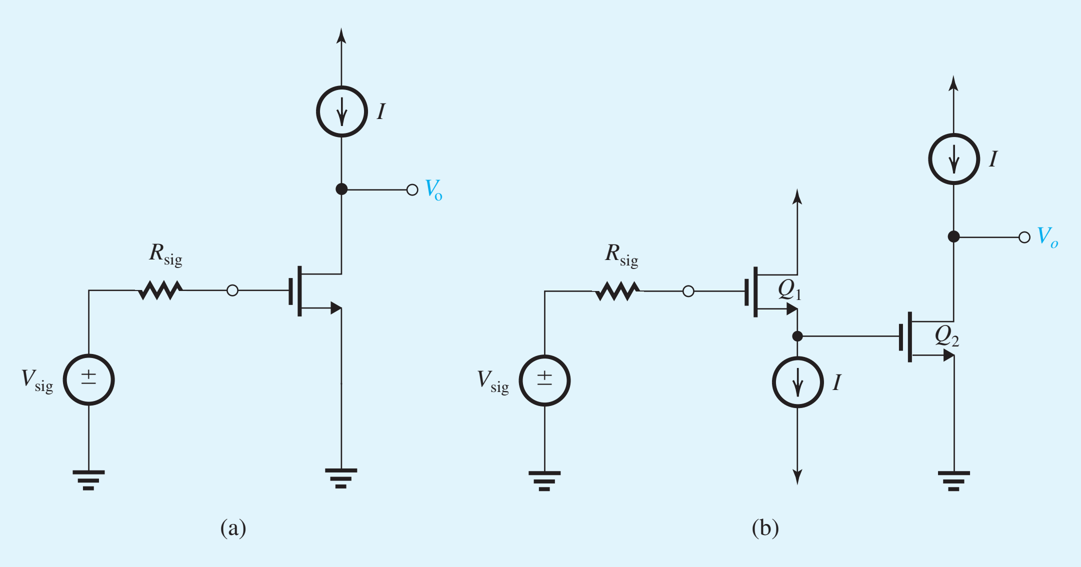


Figure P10.100

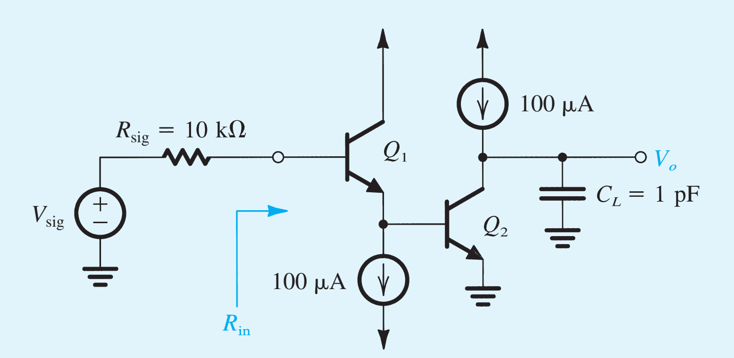


Figure P10.101

(a) Find R_{in} and the midband gain.
(b) Find an estimate of the upper 3-dB frequency f_H. Which capacitor dominates? Which one is the second most significant?

(*Hint*: Use the formulas in Example 10.13.)

10.102 Consider the circuit of Fig. P10.102 for the case: $I = 200\ \mu\text{A}$ and $V_{OV} = 0.2$ V, $R_{sig} = 100\ \text{k}\Omega$, $R_D = 50\ \text{k}\Omega$, $C_{gs} = 4$ pF, and $C_{gd} = 0.5$ pF. Find the dc gain, the high-frequency poles, and an estimate of f_H.

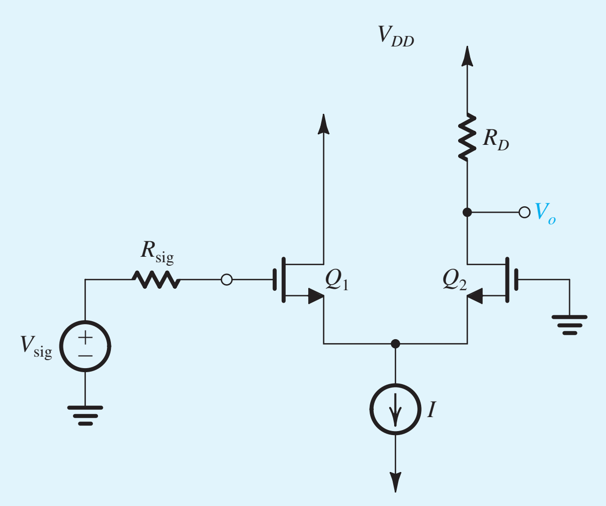


Figure P10.102

10.103 For the amplifier in Fig. 10.41(a), let $I = 1$ mA, $\beta = 120$, $f_T = 500$ MHz, and $C_\mu = 0.5$ pF, and neglect r_x and r_o. Assume that a load resistance of 10 kΩ is connected to the output terminal. If the amplifier is fed with a signal V_{sig} having a source resistance $R_{sig} = 12\ \text{k}\Omega$, find A_M and f_H.

10.104 Consider the CD–CG amplifier of Fig. 10.41(c) for the case $g_m = 5$ mA/V, $C_{gs} = 2$ pF, $C_{gd} = 0.1$ pF, C_L (at the output node) = 1 pF, and $R_{sig} = R_L = 20\ \text{k}\Omega$. Neglecting r_o, find A_M and f_H. (*Hint:* Evaluate f_H directly from the transfer function.)

D **10.105 This problem investigates the use of MOSFETs in the design of wideband amplifiers (Steininger, 1990). Such amplifiers can be realized by cascading low-gain stages.

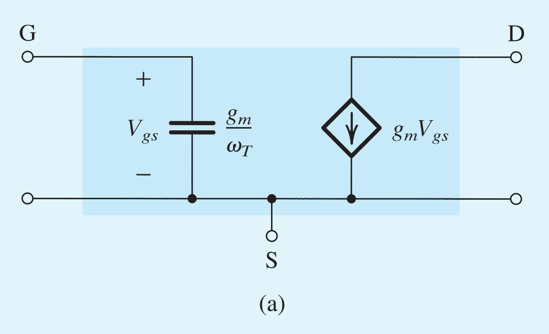
(a) Show that for the case $C_{gd} \ll C_{gs}$ and the gain of the common-source amplifier is low so that the Miller effect is negligible, the MOSFET can be modeled by the approximate equivalent circuit shown in Fig. P10.105(a), where ω_T is the unity-gain frequency of the MOSFET.

(b) Figure P10.105(b) shows an amplifier stage suitable for the realization of low gain and wide bandwidth. Transistors Q_1 and Q_2 have the same channel length L but different widths W_1 and W_2. They are biased at the same V_{GS} and have the same f_T. Use the MOSFET equivalent circuit of Fig. P10.105(a) to model this amplifier stage, assuming that its output is connected to the input of an identical stage. Show that the voltage gain V_o/V_i is given by

$$\frac{V_o}{V_i} = -\frac{G_0}{1 + \dfrac{s}{\omega_T/(G_0 + 1)}}$$

where

$$G_0 = \frac{g_{m1}}{g_{m2}} = \frac{W_1}{W_2}$$



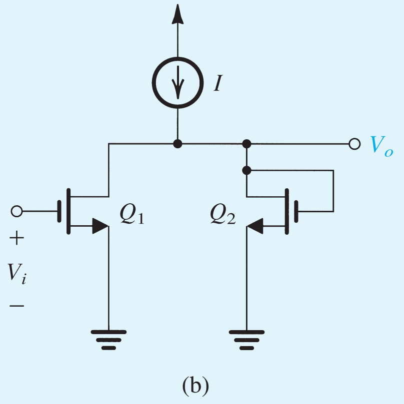


Figure P10.105

(c) For $L = 0.5\ \mu\text{m}$, $W_2 = 25\ \mu\text{m}$, $f_T = 12$ GHz, and $\mu_n C_{ox} = 200\ \mu\text{A/V}^2$, design the circuit to obtain a gain of 3 V/V per stage. Bias the MOSFETs at $V_{OV} = 0.3$ V. Specify the required values of W_1 and I. What is the 3-dB frequency achieved?

***10.106** Figure P10.106 shows an amplifier formed by cascading two CS stages. Note that the input bias voltage is not shown. Each of Q_1 and Q_2 is operated at an overdrive voltage of 0.2 V, and $|V_A| = 10$ V. The transistor capacitances are as follows: $C_{gs} = 20$ fF, $C_{gd} = 5$ fF, and $C_{db} = 5$ fF. The signal-source resistance $R_{sig} = 10$ kΩ.

(a) Find the dc voltage gain.
(b) Use the method of open-circuit time constants to determine an estimate for the 3-dB frequency f_H.

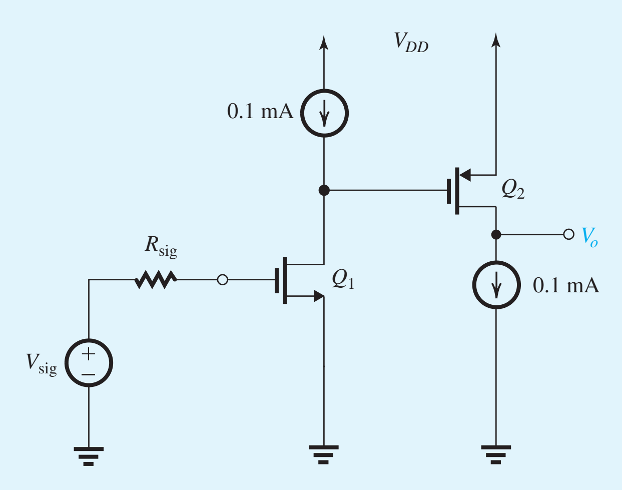


Figure P10.106

****10.107** Consider the BiCMOS amplifier shown in Fig. P10.107. The BJT has $|V_{BE}| = 0.7$ V, $\beta = 200$, $C_\mu = 0.8$ pF, and $f_T = 600$ MHz. The NMOS transistor has $V_t = 1$ V, $k_n' W/L = 2$ mA/V^2, and $C_{gs} = C_{gd} = 1$ pF.

(a) Consider the dc bias circuit. Neglect the base current of Q_2 in determining the current in Q_1. Find the dc bias currents in Q_1 and Q_2, and show that they are approximately 100 μA and 1 mA, respectively.
(b) Evaluate the small-signal parameters of Q_1 and Q_2 at their bias points.
(c) Consider the circuit at midband frequencies. First, determine the small-signal voltage gain V_o/V_i. (Note that R_G can be neglected in this process.) Then use Miller's theorem on R_G to determine the amplifier input resistance R_{in}. Finally, determine the overall voltage gain V_o/V_{sig}. Assume r_o of both transistors to be very large.
(d) Consider the circuit at low frequencies. Determine the frequency of the poles due to C_1 and C_2, and hence estimate the lower 3-dB frequency, f_L.
(e) Consider the circuit at higher frequencies. Use Miller's theorem to replace R_G with a resistance at the input. (The one at the output will be too large to matter.) Use open-circuit time constants to estimate f_H.

*****10.108** In each of the six circuits in Fig. P10.108, let $\beta = 100$, $C_\mu = 2$ pF, and $f_T = 400$ MHz, and neglect r_x and r_o. Calculate the midband gain A_M and the 3-dB frequency f_H.

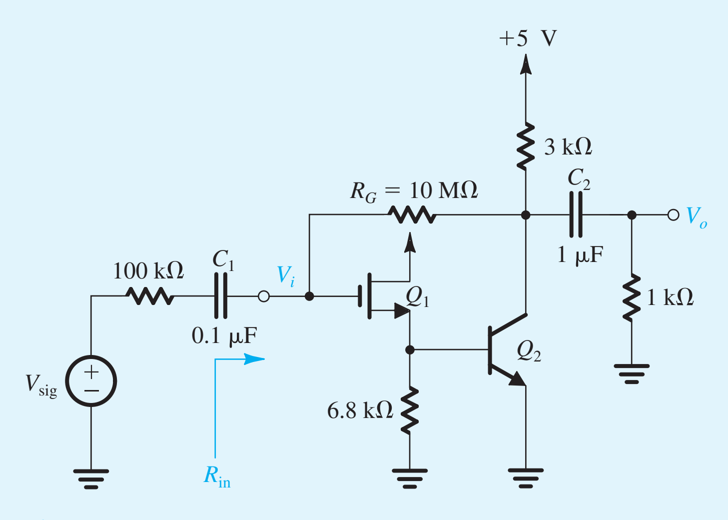


Figure P10.107

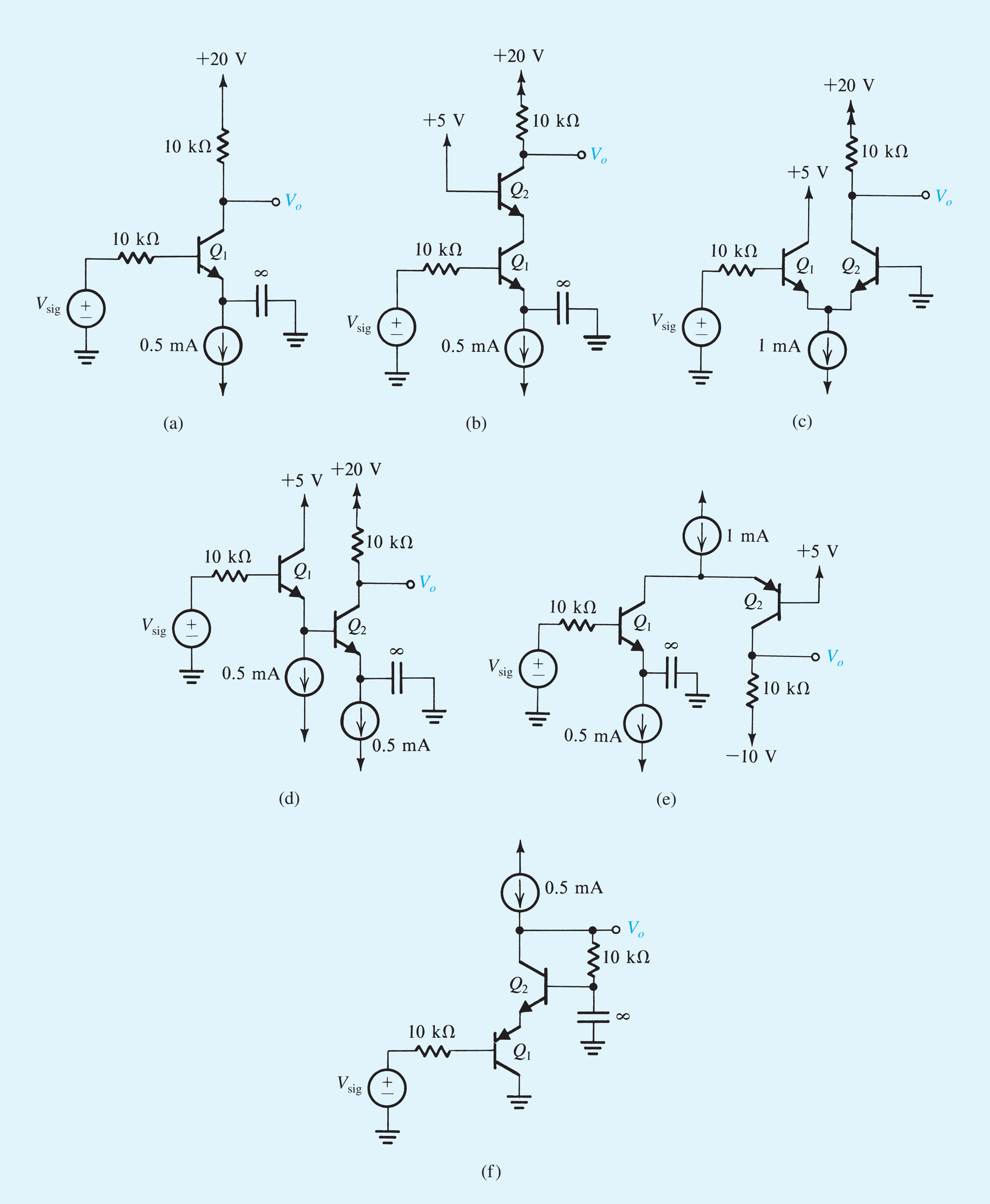


Figure P10.108

CHAPTER 11

Feedback

IN THIS CHAPTER YOU WILL LEARN

1. The general structure of the negative-feedback amplifier and the basic principle that underlies its operation.
2. The advantages of negative feedback, how these come about, and at what cost.
3. The appropriate feedback topology to employ with amplifiers of each of the four types: voltage, current, transconductance, and transresistance.
4. An intuitive and insightful approach for the analysis of practical feedback-amplifier circuits.
5. Why and how negative-feedback amplifiers can become unstable (i.e., oscillate) and how to design the circuit to ensure stable performance.

Introduction

Most physical systems incorporate some form of feedback. It is interesting to note, though, that the theory of negative feedback has been developed by electronics engineers. In his search for methods for the design of amplifiers with stable gain for use in transatlantic telephone repeaters, Harold Black, an electronics engineer with the Western Electric Company, invented the feedback amplifier in 1928. Since then, the technique has been so widely used that it is almost impossible to think of electronic circuits without some form of feedback, either implicit or explicit. Furthermore, the concept of feedback and its associated theory are currently used in areas other than engineering, such as in the modeling of biological systems.

Feedback can be either **negative** or **positive**. In amplifier design, negative feedback is applied to effect one or more of the following goals:

1. *Desensitize the gain:* that is, make the value of the gain less sensitive to variations in the values of circuit components, such as might be caused by changes in temperature.
2. *Reduce nonlinear distortion:* that is, make the output proportional to the input (in other words, make the gain constant, independent of signal level).
3. *Reduce the effect of noise:* that is, minimize the contribution to the output of unwanted electric signals generated, either by the circuit components themselves or by extraneous interference.
4. *Control the input and output resistances:* that is, raise or lower the input and output resistances by the selection of an appropriate feedback topology.
5. *Extend the bandwidth* of the amplifier.

All of the desirable properties above are obtained at the expense of a reduction in gain. It will be shown that the gain-reduction factor, called the **amount of feedback**, is the factor by which the circuit is desensitized, by which the input resistance of a voltage amplifier is increased, by which the bandwidth is extended, and so on. In short, *the basic idea of negative feedback is to trade off gain for other desirable properties*. This chapter is devoted to the study of negative-feedback amplifiers: their analysis, design, and characteristics.

Under certain conditions, the negative feedback in an amplifier can become positive and of such a magnitude as to cause oscillation. In fact, in Chapter 18 we will study the use of positive feedback in the design of oscillators and bistable circuits. Here, in this chapter, however, we are interested in the design of stable amplifiers. We shall therefore study the stability problem of negative-feedback amplifiers and their potential for oscillation.

It should not be implied, however, that positive feedback always leads to instability. In fact, positive feedback is quite useful in a number of nonregenerative applications, such as the design of active filters, which are studied in Chapter 17.

Before we begin our study of negative feedback, we wish to remind the reader that we have already encountered negative feedback in a number of applications. Almost all op-amp circuits (Chapter 2) employ negative feedback. Another popular application of negative feedback is the use of the emitter resistance R_E to stabilize the bias point of bipolar transistors and to increase the input resistance, bandwidth, and linearity of a BJT amplifier. In addition, the source follower and the emitter follower both employ a large amount of negative feedback. The question then arises about the need for a formal study of negative feedback. As will be appreciated by the end of this chapter, the formal study of feedback provides an invaluable tool for the analysis and design of electronic circuits. Also, the insight gained by thinking in terms of feedback can be extremely profitable.

11.1 The General Feedback Structure

11.1.1 Signal-Flow Diagram

Figure 11.1 shows the basic structure of a feedback amplifier. Rather than showing voltages and currents, Fig. 11.1 is a signal-flow diagram, where each of the quantities x can represent either a voltage or a current signal. The basic amplifier is unilateral and has a gain A, known as the **open-loop gain**; thus its output x_o is related to the input x_i by

$$x_o = Ax_i \tag{11.1}$$

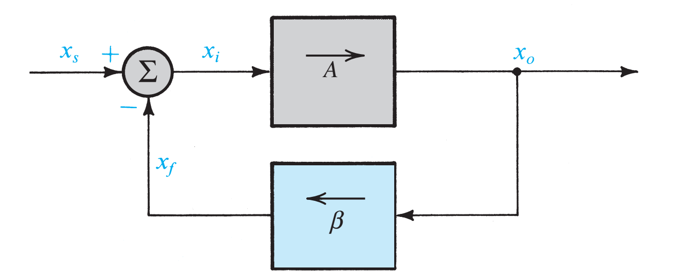


Figure 11.1 General structure of the feedback amplifier. This is a signal-flow diagram, and the quantities x represent either voltage or current signals.

The *feedback network* measures or samples the output signal x_o and provides a *feedback signal* x_f that is related to x_o by the **feedback factor** β,

$$x_f = \beta x_o \tag{11.2}$$

It is assumed that connecting the feedback network to the amplifier output does not change the gain A or the value of x_o; that is, *the feedback network does not load the amplifier output*. Also, the feedback network is unilateral.

The feedback signal x_f is *subtracted* from the source signal x_s, which is the input to the complete feedback amplifier,[1] to produce the signal x_i, which is the input to the basic amplifier,

$$x_i = x_s - x_f \tag{11.3}$$

Here we note that it is this subtraction that makes the feedback negative. In essence, *negative feedback reduces the signal that appears at the input of the basic amplifier*. Here, too, we assume that connecting the output of the feedback network to the amplifier input, through the subtractor or differencing circuit, does not change the gain A; that is, *the feedback network does not load the amplifier input.*

11.1.2 The Closed-Loop Gain

The gain of the feedback amplifier, known as the closed-loop gain or the **gain-with-feedback** and denoted A_f, is defined as

$$A_f \equiv \frac{x_o}{x_s}$$

Combining Eqs. (11.1) through (11.3) provides the following expression for A_f:

$$A_f = \frac{A}{1 + A\beta} \tag{11.4}$$

The quantity $A\beta$ is called the **loop gain**, a name that follows from Fig. 11.1. For the feedback to be negative, the loop gain $A\beta$ must be positive; that is, the feedback signal x_f should have the same sign as x_s, thus resulting in a smaller difference signal x_i. Equation (11.4) indicates that for positive $A\beta$ the gain with feedback A_f will be smaller than the open-loop gain A by a factor equal to $1 + A\beta$, which is called the **amount of feedback**.

If, as is the case in many circuits, the loop gain $A\beta$ is large, $A\beta \gg 1$, then from Eq. (11.4) it follows that

$$A_f \simeq \frac{1}{\beta} \tag{11.5}$$

which is a very interesting result: *When the loop gain is large, the gain of the feedback amplifier is almost entirely determined by the feedback network*. Since the feedback network usually consists of passive components, which usually can be chosen to be as accurate as one wishes, the advantage of negative feedback in obtaining accurate, predictable, and stable gain

[1] In earlier chapters, we used the subscript "sig" for quantities associated with the signal source (e.g., v_{sig} and R_{sig}). We did that to avoid confusion with the subscript "*s*," which is usually used with FETs to denote quantities associated with the source terminal of the transistor. At this point, however, it is expected that readers have become sufficiently familiar with the subject that the possibility of confusion is minimal. Therefore, we will revert to using the simpler subscript *s* for signal-source quantities.

should be apparent. In other words, the overall gain will have very little dependence on the gain of the basic amplifier, A, a desirable property because the gain A is usually a function of many manufacturing and application parameters, some of which might have wide tolerances. We have seen a dramatic illustration of all of these effects in op-amp circuits in Chapter 2, where the closed-loop gain is almost entirely determined by the feedback elements. Generally, we will consider $(1/\beta)$ to be the ideal value of A_f.

Equations (11.1) through (11.3) can be combined to obtain the following expression for the feedback signal x_f:

$$x_f = \frac{A\beta}{1+A\beta}x_s \tag{11.6}$$

Thus for $A\beta \gg 1$ we see that $x_f \simeq x_s$, which implies that the signal x_i at the input of the basic amplifier is reduced to almost zero. Thus if a large amount of negative feedback is employed, the feedback signal x_f becomes an almost identical replica of the input signal x_s. The difference between x_s and x_f, which is x_i, is sometimes referred to as the **error signal**.[2] Accordingly, the **input differencing circuit** is often also called a **comparison circuit**. (It is also known as a **mixer**.) An expression for x_i can be easily determined as

$$x_i = \frac{1}{1+A\beta}x_s \tag{11.7}$$

from which we can verify that for $A\beta \gg 1$, x_i becomes very small. An outcome of this property is the tracking of the two input terminals of an op amp. Observe that negative feedback reduces the signal that appears at the input terminals of the basic amplifier by the amount of feedback $(1 + A\beta)$. As will be seen later, it is this reduction of input signal that results in the increased linearity of the feedback amplifier.[3]

11.1.3 The Loop Gain

From the discussion above we see that the loop gain $A\beta$ is a very important—in fact, the most important—characteristic parameter of a feedback amplifier:

1. The sign of $A\beta$ determines the polarity of the feedback; the loop gain $A\beta$ must be positive for the feedback to be negative.
2. The magnitude of $A\beta$ determines how close the closed-loop gain A_f is to the ideal value of $1/\beta$.
3. The magnitude of $A\beta$ determines the amount of feedback $(1+A\beta)$ and hence, as we shall see in the next section, the magnitude of the various improvements in amplifier performance resulting from the negative feedback.
4. As we shall see in later sections, the inevitable variation of $A\beta$ with frequency can cause $A\beta$ to become negative, which in turn can cause the feedback amplifier to become unstable. It follows that the design of a stable feedback amplifier may involve modifying the frequency behaviors of its loop gain $A\beta$ appropriately (Section 11.10).

[2] This terminology is more common in feedback control systems than in feedback amplifiers.

[3] We have in fact already seen examples of this: adding a resistance R_e in the emitter of a CE amplifier (or a resistance R_s in the source of a CS amplifier) increases the linearity of these amplifiers because for the same input signal as before, v_{be} and v_{gs} are now smaller (by the amount of feedback).

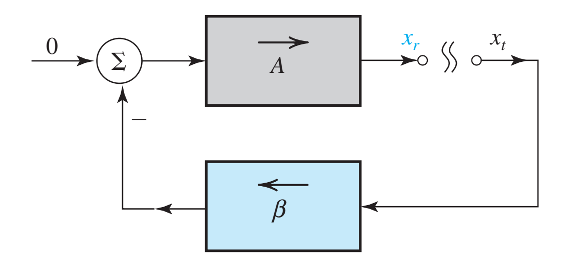


Figure 11.2 Determining the loop gain by breaking the feedback loop at the output of the basic amplifier, applying a test signal x_t, and measuring the returned signal x_r : $A\beta \equiv -x_r/x_t$.

The significance of the loop gain requires us to consider its determination. Reference to Fig. 11.1 indicates that the value of the loop gain $A\beta$ can be determined as follows:

1. Set $x_s = 0$.
2. Break the feedback loop at a convenient location, ensuring that the values of A and β *do not change*. Since we assumed that the feedback network does not load the amplifier output, we can break the loop at the amplifier output (see Fig. 11.2) without causing A to change.
3. Apply a test signal x_t to the input of the loop (where the break has been made) and determine the *returned signal* x_r at the loop output (i.e., at the other side of the break). From Fig. 11.2 we see that

$$x_r = -A\beta x_t$$

and the loop gain $A\beta$ is obtained as

$$A\beta = -\frac{x_r}{x_t} \tag{11.8}$$

We observe that since $A\beta$ is positive, the returned signal x_r will be out of phase with the test signal x_t, verifying that the feedback is indeed negative. In fact, this approach is used qualitatively to ascertain the polarity of the feedback. We will have a lot more to say about the loop gain in subsequent sections.

Example 11.1

The noninverting op-amp configuration shown in Fig. 11.3(a) provides a direct implementation of the feedback loop of Fig. 11.1.

(a) Assume that the op amp has infinite input resistance and zero output resistance. Find an expression for the feedback factor β.

(b) Find the condition the open-loop gain A must satisfy so that the closed-loop gain A_f is almost entirely determined by the feedback network. Also, give the value of A_f in this case.

Example 11.1 *continued*

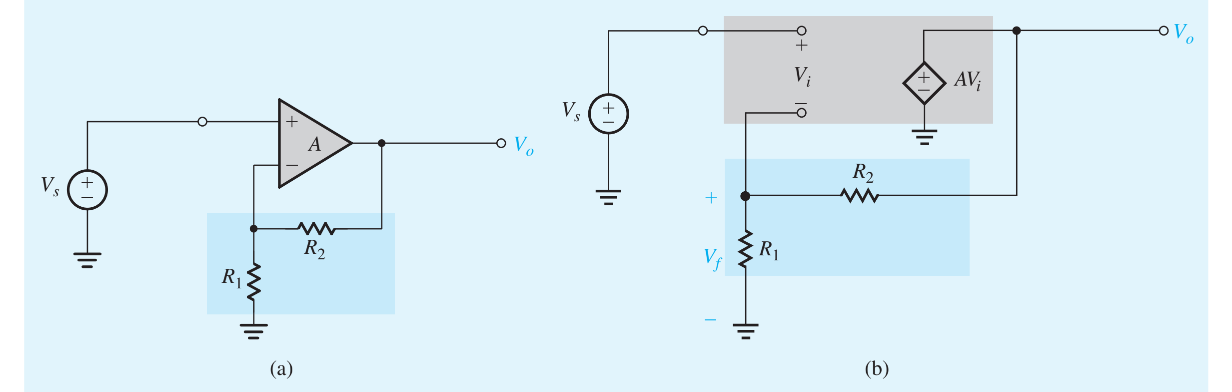


Figure 11.3 **(a)** A noninverting op-amp circuit for Example 11.1. **(b)** The circuit in **(a)** with the op amp replaced with its equivalent circuit.

(c) If the open-loop gain $A = 10^4$ V/V, find R_2/R_1 to obtain a closed-loop gain A_f of 10 V/V.
(d) What is the amount of feedback in decibels?
(e) If $V_s = 1$ V, find V_o, V_f, and V_i.
(f) If A decreases by 20%, what is the corresponding decrease in A_f?

Solution

(a) To be able to see more clearly the direct correspondence between the circuit in Fig. 11.3(a) and the block diagram in Fig. 11.1, we replace the op amp with its equivalent-circuit model, as shown in Fig. 11.3(b). Since the op amp is assumed to have infinite input resistance and zero output resistance, its model is simply an ideal voltage-controlled voltage source of gain A. From Fig. 11.3(b) we observe that the feedback network, consisting of the voltage divider (R_1, R_2), is connected directly to the output and feeds a signal V_f to the inverting input terminal of the op amp. It is important at this point to note that the zero output resistance of the op amp causes the output voltage to be AV_i irrespective of the values of R_1 and R_2. That is what we meant by the statement that in the block diagram of Fig. 11.1, the feedback network is assumed to not load the basic amplifier. Now we can easily determine the feedback factor β from

$$\beta \equiv \frac{V_f}{V_o} = \frac{R_1}{R_1 + R_2}$$

Let's next examine how V_f is subtracted from V_s at the input side. The subtraction is effectively performed by the differential action of the op amp; by its very nature, a differential-input amplifier takes the difference between the signals at its two input terminals. Observe also that because the input resistance of the op amp is assumed to be infinite, no current flows into the negative input terminal of the op amp and that the feedback network does not load the amplifier at the input side.

(b) The closed-loop gain A_f is given by

$$A_f = \frac{A}{1 + A\beta}$$

To make A_f nearly independent of A, we must ensure that the loop gain $A\beta$ is much larger than unity

$$A\beta \gg 1$$

in which case

$$A_f \simeq 1/\beta$$

Thus,

$$A/A_f \gg 1$$

or equivalently,

$$A \gg A_f$$

and

$$A_f \simeq \frac{1}{\beta} = 1 + \frac{R_2}{R_1}$$

(c) For $A = 10^4$ V/V and $A_f = 10$ V/V, we see that $A \gg A_f$, thus we can select R_1 and R_2 to obtain

$$\beta \simeq \frac{1}{A_f} = 0.1$$

Thus,

$$\frac{1}{\beta} = 1 + \frac{R_2}{R_1} = A_f = 10$$

which yields

$$R_2/R_1 = 9$$

A more exact value for the required ratio R_2/R_1 can be obtained from

$$A_f = \frac{A}{1 + A\beta}$$

$$10 = \frac{10^4}{1 + 10^4\beta}$$

which results in

$$\beta = 0.0999$$

and,

$$\frac{R_2}{R_1} = 9.01$$

Example 11.1 *continued*

(d) The amount of feedback is

$$1 + A\beta = \frac{A}{A_f} = \frac{10^4}{10} = 1000$$

which is 60 dB.
(e) For $V_s = 1$ V,

$$V_o = A_f V_s = 10 \times 1 = 10 \text{ V}$$
$$V_f = \beta V_o = 0.0999 \times 10 = 0.999 \text{ V}$$
$$V_i = \frac{V_o}{A} = \frac{10}{10^4} = 0.001 \text{ V}$$

Note that if we had used the approximate value of $\beta = 0.1$, we would have obtained $V_f = 1$ V and $V_i = 0$ V.
(f) If A decreases by 20%, thus becoming

$$A = 0.8 \times 10^4 \text{ V/V}$$

the value of A_f becomes

$$A_f = \frac{0.8 \times 10^4}{1 + 0.8 \times 10^4 \times 0.0999} = 9.9975 \text{ V/V}$$

that is, it decreases by 0.025%, which is less than the percentage change in A by approximately a factor $(1 + A\beta)$.

EXERCISES

11.1 Repeat Example 11.1 (c) to (f) for $A = 100$ V/V.
Ans. (c) 10.11; (d) 20 dB; (e) 10 V, 0.9 V, 0.1 V; (f) 2.44%

11.2 Repeat Example 11.1 (c) to (f) for $A_f = 10^3$ V/V. For (e) use $V_s = 0.01$ V.
Ans. (c) 1110.1; (d) 20 dB; (e) 10 V, 0.009 V, 0.001 V; (f) 2.44%

11.1.4 Summary

We conclude this section by presenting in Table 11.1 a summary of the important parameters and formulas that characterize the ideal negative-feedback amplifier structure of Fig. 11.1.

Table 11.1 Summary of the Parameters and Formulas for the Ideal Feedback-Amplifier Structure of Fig. 11.1

- Open-loop gain $\equiv A$
- Feedback factor $\equiv \beta$
- Loop gain $\equiv A\beta$ (positive number)
- Amount of feedback $\equiv 1 + A\beta$
- Closed-loop gain $\equiv A_f = \dfrac{x_o}{x_s} = \dfrac{A}{1 + A\beta}$
- Feedback signal $\equiv x_f = \dfrac{A\beta}{1 + A\beta} x_s$
- Input signal to basic amplifier $\equiv x_i = \dfrac{1}{1 + A\beta} x_s$
- Closed-loop gain as a function of the ideal value $\dfrac{1}{\beta}$: $A_f = \left(\dfrac{1}{\beta}\right) \dfrac{1}{1 + 1/A\beta}$
- For large loop gain, $A\beta \gg 1$,

$A_f \simeq \dfrac{1}{\beta}$ $\quad x_f \simeq x_s$ $\quad x_i \simeq 0$

11.2 Some Properties of Negative Feedback

The properties of negative feedback were mentioned in the introduction. In the following, we shall consider some of these properties in more detail.

11.2.1 Gain Desensitivity

The effect of negative feedback on desensitizing the closed-loop gain was demonstrated in Example 11.1, where we saw that a 20% reduction in the gain of the basic amplifier gave rise to only a 0.025% reduction in the gain of the closed-loop amplifier. This sensitivity-reduction property can be analytically established as follows.

Assume that β is constant. Taking differentials of both sides of Eq. (11.4) results in

$$dA_f = \frac{dA}{(1 + A\beta)^2} \tag{11.9}$$

Dividing Eq. (11.9) by Eq. (11.4) yields

$$\frac{dA_f}{A_f} = \frac{1}{(1 + A\beta)} \frac{dA}{A} \tag{11.10}$$

which says that the percentage change in A_f (due to variations in some circuit parameter) is smaller than the percentage change in A by a factor equal to the amount of feedback. For this reason, the amount of feedback, $1 + A\beta$, is also known as the **desensitivity factor**.

EXERCISE

11.3 An amplifier with a nominal gain $A = 1000$ V/V exhibits a gain change of 10% as the operating temperature changes from 25°C to 75°C. If it is required to constrain the change to 0.1% by applying negative feedback, what is the largest closed-loop gain possible? If three of these feedback amplifiers are placed in cascade, what overall gain and gain stability are achieved?
Ans. 10 V/V; 1000 V/V, with a maximum variability of 0.3% over the specified temperature range.

11.2.2 Bandwidth Extension

Consider an amplifier whose high-frequency response is characterized by a single pole. Its gain at mid and high frequencies can be expressed as

$$A(s) = \frac{A_M}{1 + s/\omega_H} \tag{11.11}$$

where A_M denotes the midband gain and ω_H is the upper 3-dB frequency. Application of negative feedback, with a frequency-independent factor β, around this amplifier results in a closed-loop gain $A_f(s)$ given by

$$A_f(s) = \frac{A(s)}{1 + \beta A(s)}$$

Substituting for $A(s)$ from Eq. (11.11) results, after a little manipulation, in

$$A_f(s) = \frac{A_M/(1 + A_M\beta)}{1 + s/\omega_H(1 + A_M\beta)} \tag{11.12}$$

Thus the feedback amplifier will have a midband gain of $A_M/(1 + A_M\beta)$ and an upper 3-dB frequency ω_{Hf} given by

$$\omega_{Hf} = \omega_H(1 + A_M\beta) \tag{11.13}$$

It follows that the upper 3-dB frequency is increased by a factor equal to the amount of feedback.

Similarly, it can be shown that if the open-loop gain is characterized by a dominant low-frequency pole giving rise to a lower 3-dB frequency ω_L, then the feedback amplifier will have a lower 3-dB frequency ω_{Lf},

$$\omega_{Lf} = \frac{\omega_L}{1 + A_M\beta} \tag{11.14}$$

Note that the amplifier bandwidth is increased by the same factor by which its midband gain is decreased, *maintaining the gain–bandwidth product at a constant value*. This point is further illustrated by the Bode plot in Fig. 11.4.

Finally, note that the action of negative feedback in extending the amplifier bandwidth should not be surprising: Negative feedback works to minimize the change in gain magnitude, including its change with frequency.

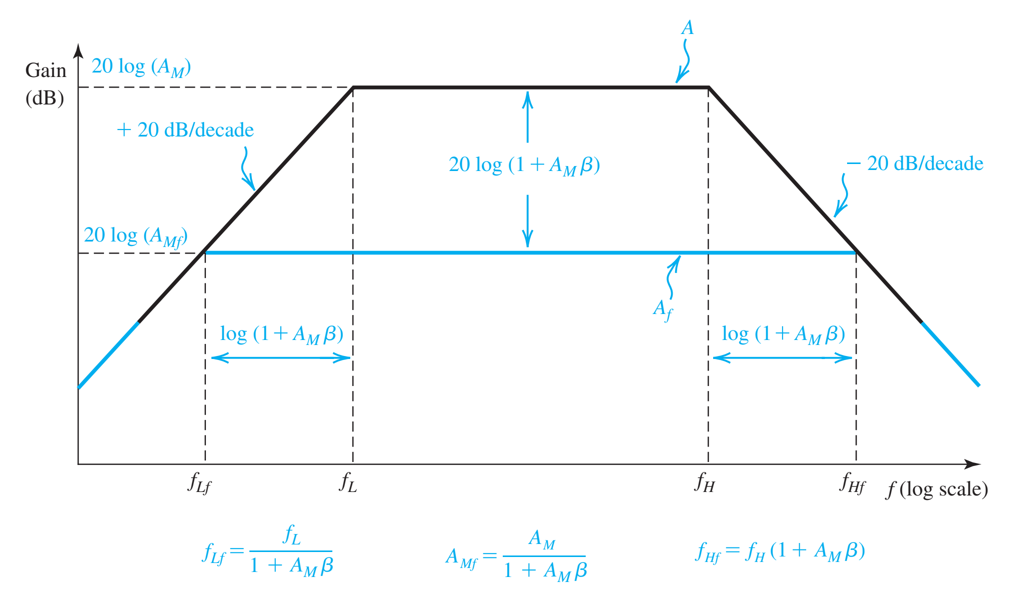


Figure 11.4 Application of negative feedback reduces the midband gain, increases f_H, and reduces f_L, all by the same factor, $(1 + A_M\beta)$, which is equal to the amount of feedback.

EXERCISE

11.4 Consider the noninverting op-amp circuit of Example 11.1. Let the open-loop gain A have a low-frequency value of 10^4 and a uniform –6-dB/octave rolloff at high frequencies with a 3-dB frequency of 100 Hz. Find the low-frequency gain and the upper 3-dB frequency of a closed-loop amplifier with $R_1 = 1\text{ k}\Omega$ and $R_2 = 9\text{ k}\Omega$.
Ans. 9.99 V/V; 100.1 kHz

11.2.3 Interference Reduction

Negative feedback can be employed to reduce the interference in an amplifier or, more precisely, to increase the ratio of signal to interference. However, as we shall now explain, this interference-reduction process is possible only under certain conditions. Consider the situation illustrated in Fig. 11.5. Figure 11.5(a) shows an amplifier with gain A_1, an input signal V_s, and interference, V_n. It is assumed that for some reason this amplifier suffers from interference and that the interference can be assumed to be introduced at the input of the amplifier. The **signal-to-interference ratio** for this amplifier is

$$S/I = V_s/V_n \tag{11.15}$$

Consider next the circuit in Fig. 11.5(b). Here we assume that it is possible to build another amplifier stage with gain A_2 that does not suffer from the interference problem. If this is the

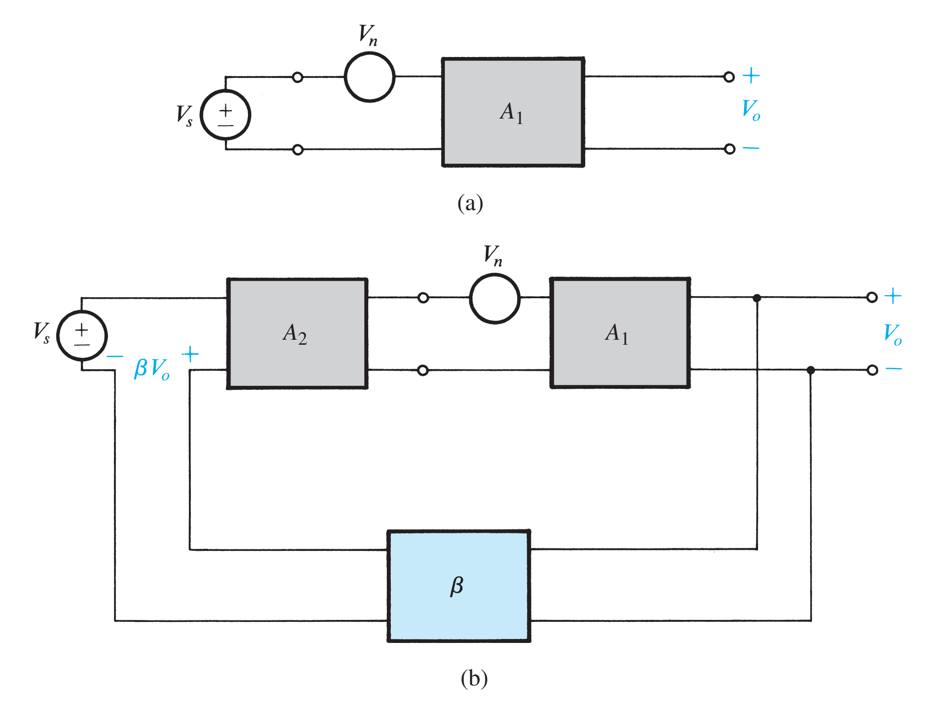


Figure 11.5 Illustrating the application of negative feedback to improve the signal-to-interference ratio in amplifiers.

case, then we may precede our original amplifier A_1 by the *clean* amplifier A_2 and apply negative feedback around the overall cascade of such an amount as to keep the overall gain unchanged. The output voltage of the circuit in Fig. 11.5(b) can be found by superposition:

$$V_o = V_s \frac{A_1 A_2}{1 + A_1 A_2 \beta} + V_n \frac{A_1}{1 + A_1 A_2 \beta} \tag{11.16}$$

Thus the signal-to-interference ratio at the output becomes

$$\frac{S}{I} = \frac{V_s}{V_n} A_2 \tag{11.17}$$

which is A_2 times higher than in the original case.

We emphasize once more that the improvement in signal-to-interference ratio by the application of feedback is possible only if one can precede the interference-prone stage by a (relatively) interference-free stage. This situation, however, is not uncommon in practice. The best example is found in the output power-amplifier stage of an audio amplifier. Such a stage usually suffers from a problem known as **power-supply hum**. The problem arises because of the large currents that this stage draws from the power supply and because it is difficult to provide adequate power-supply filtering inexpensively. The power-output stage is required to provide large power gain but little or no voltage gain. We may therefore precede the power-output stage by a small-signal amplifier that provides large voltage gain and then apply a large amount of negative feedback, thus restoring the voltage gain to its original value. Since the small-signal amplifier can be fed from another, less hefty (and hence better regulated) power supply, it will not suffer from the hum problem. The

hum at the output will then be reduced by the amount of the voltage gain of this added **preamplifier**.

EXERCISE

11.5 Consider a power-output stage with voltage gain $A_1 = 1$, an input signal $V_s = 1$ V, and a hum V_n of 1 V. Assume that this power stage is preceded by a small-signal stage with gain $A_2 = 100$ V/V and that overall feedback with $\beta = 1$ is applied. If V_s and V_n remain unchanged, find the signal and interference voltages at the output and hence the improvement in S/I.
Ans. $\simeq$1 V; $\simeq$0.01 V; 100 (40 dB)

11.2.4 Reduction in Nonlinear Distortion

Curve (a) in Fig. 11.6 shows the transfer characteristic v_O versus v_I of an amplifier. As indicated, the characteristic is piecewise linear, with the voltage gain changing from 1000 to 100 and then to 0. This nonlinear transfer characteristic will result in this amplifier generating a large amount of nonlinear distortion.

The amplifier transfer characteristic can be considerably **linearized** (i.e., made less nonlinear) through the application of negative feedback. That this is possible should not be too surprising, since we have already seen that negative feedback reduces the dependence of the overall closed-loop amplifier gain on the open-loop gain of the basic amplifier. Thus large

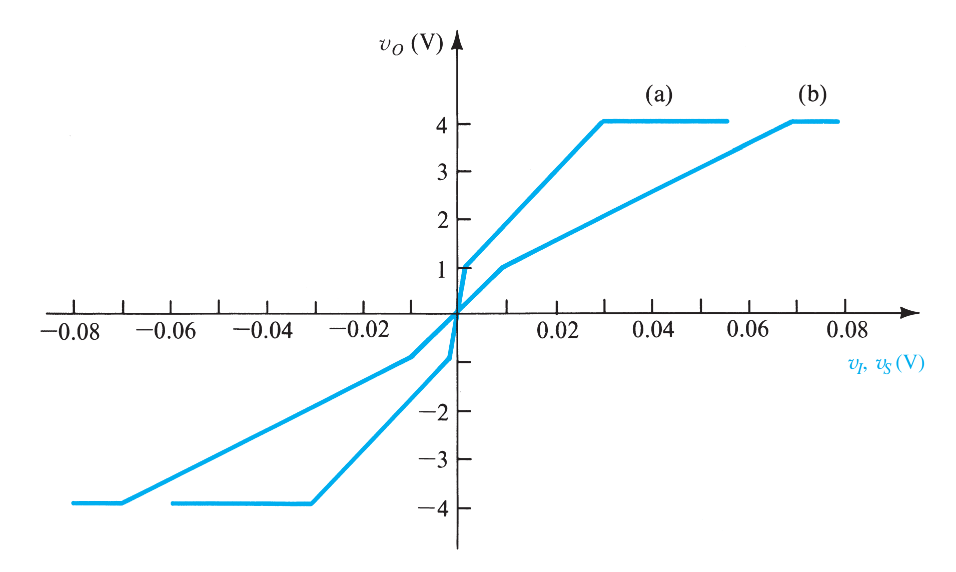


Figure 11.6 Illustrating the application of negative feedback to reduce the nonlinear distortion in amplifiers. Curve (a) shows the amplifier transfer characteristic (v_O versus v_I) without feedback. Curve (b) shows the characteristic (v_O versus v_S) with negative feedback ($\beta = 0.01$) applied.

changes in open-loop gain (1000 to 100 in this case) give rise to much smaller corresponding changes in the closed-loop gain.

To illustrate, let us apply negative feedback with $\beta = 0.01$ to the amplifier whose open-loop voltage transfer characteristic is depicted in Fig. 11.6. The resulting transfer characteristic of the closed-loop amplifier, v_O versus v_S, is shown in Fig. 11.6 as curve (b). Here the slope of the steepest segment is given by

$$A_{f1} = \frac{1000}{1 + 1000 \times 0.01} = 90.9$$

and the slope of the next segment is given by

$$A_{f2} = \frac{100}{1 + 100 \times 0.01} = 50$$

Thus the order-of-magnitude change in slope has been considerably reduced. The price paid, of course, is a reduction in voltage gain. Thus if the overall gain has to be restored, a preamplifier should be added. This preamplifier should not present a severe nonlinear-distortion problem, since it will be dealing with smaller signals.

Finally, it should be noted that negative feedback can do nothing at all about amplifier saturation, since in saturation the gain is very small (almost zero) and hence the amount of feedback is almost unity.

11.3 The Feedback Voltage Amplifier

Based on the quantity to be amplified (voltage or current) and on the desired form of output (voltage or current), amplifiers can be classified into four categories. These categories were discussed in Chapter 1. In this section we study the most common amplifier type: the voltage amplifier. We begin by identifying the appropriate configuration for applying negative feedback to a voltage amplifier. Then, we present a simple method for the analysis of the feedback voltage amplifier. The method makes use of the loop gain $A\beta$, whose determination was discussed in Section 11.1.3.

11.3.1 The Series–Shunt Feedback Topology

Voltage amplifiers are intended to amplify an input voltage signal and provide an output voltage signal. The voltage amplifier is essentially a voltage-controlled voltage source. The input resistance is required to be high, and the output resistance is required to be low. Since the signal source is essentially a voltage source, it is appropriately represented in terms of a Thévenin equivalent circuit. As the output quantity of interest is the output voltage, the feedback network should *sample* the output *voltage*, just as a voltmeter measures a voltage. Also, because of the Thévenin representation of the source, the feedback signal x_f should be a *voltage* that can be *mixed* with the source voltage in *series*.

From the discussion above, it follows that the most suitable feedback topology for the voltage amplifier is the **voltage-mixing**, **voltage-sampling** one shown in Fig. 11.7. Because of the series connection at the input and the parallel or shunt connection at the output, this feedback topology is also known as **series–shunt feedback**. As will be shown, this topology not only stabilizes the voltage gain V_o/V_s but also results in a higher input resistance R_{in} (intuitively, a result of the series connection at the input) and a lower output resistance R_{out}

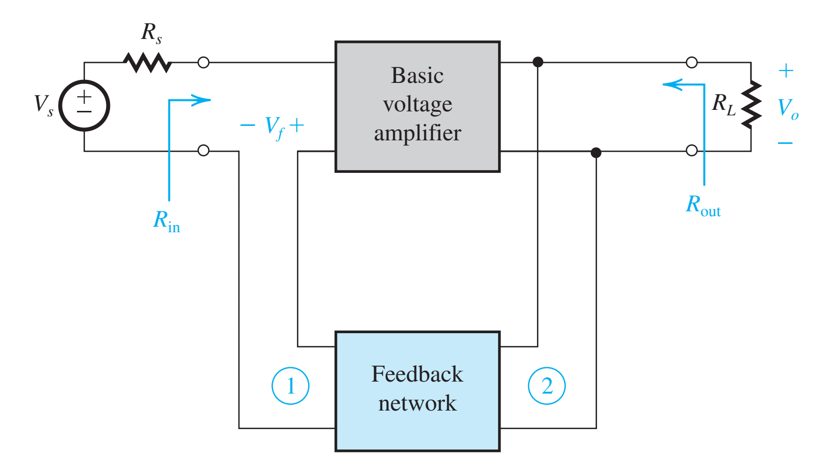


Figure 11.7 Block diagram of a feedback voltage amplifier. Here the appropriate feedback topology is series–shunt.

(intuitively, a result of the parallel connection at the output), which are desirable properties for a voltage amplifier.

The increased input resistance results because V_f subtracts from V_s, resulting in a smaller signal V_i at the input of the basic amplifier. The lower V_i, in turn, causes the input current to be smaller, with the result that the resistance seen by V_s will be larger. We shall derive a formula for the input resistance of the feedback voltage amplifier in the next section.

The decreased output resistance results because the feedback works to keep V_o as constant as possible. Thus if the current drawn from the amplifier output changes by ΔI_o, the change ΔV_o in V_o will be lower than it would have been if feedback were not present. Thus the output resistance $\Delta V_o/\Delta I_o$ will be lower than that of the open-loop amplifier. In the following section we shall derive an expression for the output resistance of the feedback voltage amplifier.

11.3.2 Examples of Series–Shunt Feedback Amplifiers

Three examples of series–shunt feedback amplifiers are shown in Fig. 11.8. The amplifier in Fig. 11.8(a) is the familiar noninverting op-amp configuration. The feedback network, composed of the voltage divider (R_1, R_2), develops a voltage V_f that is applied to the negative input terminal of the op amp. The subtraction of V_f from V_s is achieved by utilizing the differencing action of the op-amp differential input. For the feedback to be negative, V_f must be of the same polarity as V_s, thus resulting in a smaller signal at the input of the basic amplifier. To ascertain that this is the case, we follow the signal around the loop, as follows: As V_s increases, V_o increases and the voltage divider causes V_f to increase. Thus the change in V_f is of the same polarity as the change in V_s, and the feedback is negative.

The second feedback voltage amplifier, shown in Fig. 11.8(b), utilizes two MOSFET amplifier stages in cascade. The output voltage V_o is sampled by the feedback network composed of the voltage divider (R_1, R_2), and the feedback signal V_f is fed to the source terminal of Q_1. The subtraction is implemented by applying V_s to the gate of Q_1 and V_f to its source, with the result that the signal at this amplifier input $V_i = V_{gs} = V_s - V_f$. To ascertain that the feedback is negative, let V_s increase. The drain voltage of Q_1 will decrease, and since this is applied to the gate of Q_2, its drain voltage V_o will increase. The feedback network will

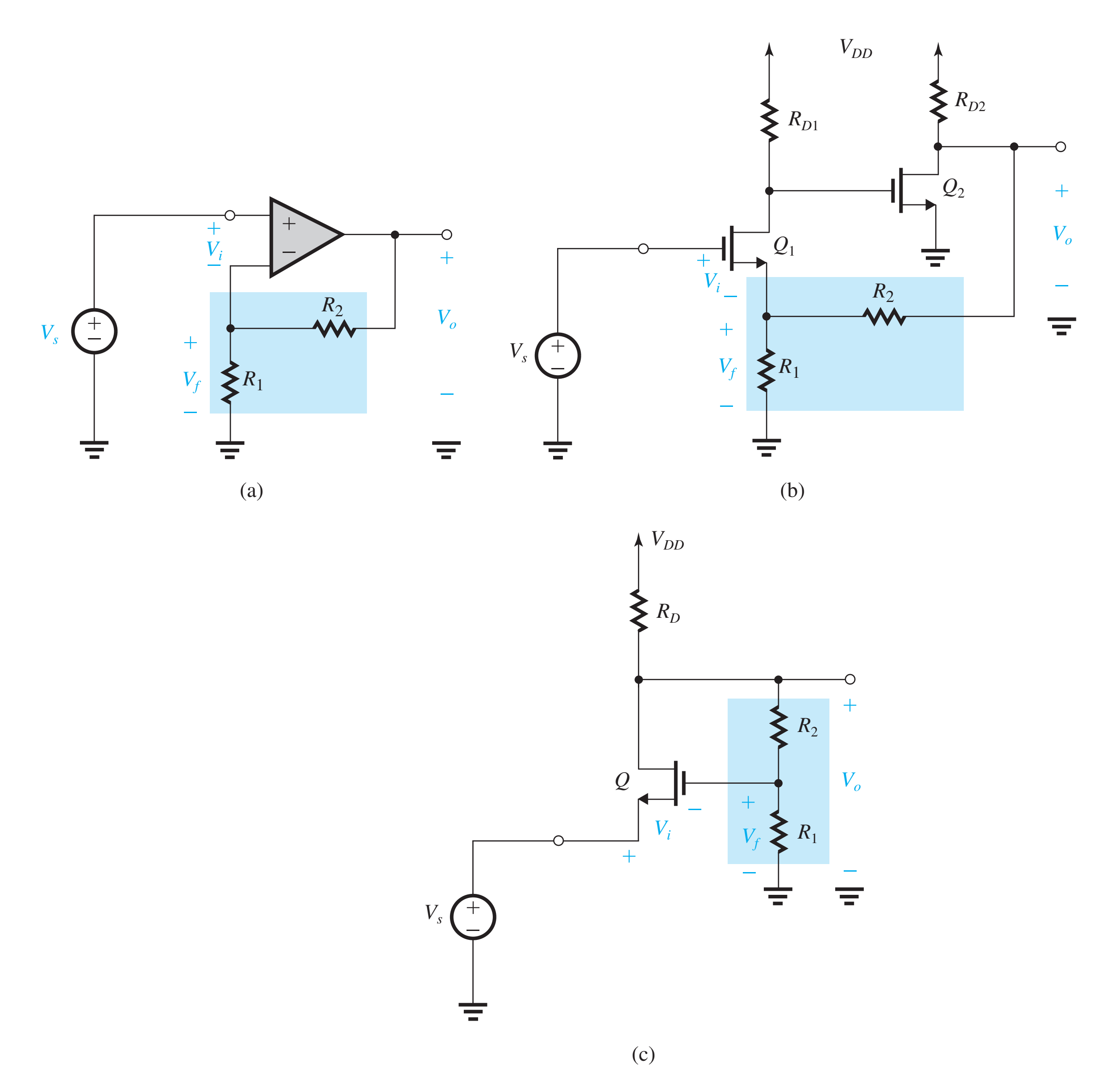


Figure 11.8 Examples of a feedback voltage amplifier. All these circuits employ series–shunt feedback. Note that the dc bias circuits are only partially shown.

then cause V_f to increase, which is the same polarity initially assumed for the change in V_s. Thus the feedback is indeed negative.

The third example of series–shunt feedback, shown in Fig. 11.8(c), utilizes a CG transistor Q with a fraction V_f of the output voltage V_o fed back to the gate through a voltage divider (R_1, R_2). Observe that the subtraction of V_f from V_s is effected by applying V_s to the source, thus the input V_i to the CG amplifier is obtained as $V_s - V_f$. As usual, however, we must check the polarity of the feedback: If V_s increases, V_d (which is V_o) will increase and V_f will correspondingly increase. Thus V_f and V_s change in the same direction, verifying that the feedback is negative.

FEEDBACK—HISTORICAL NOTE:

The idea of feedback as an element of self-regulating behavior dates back to the eighteenth century, but the term itself did not appear in the context of a discussion on economics until the 1860s. Still later, in 1909, Karl Ferdinand Braun, a German physicist working at the University of Strasbourg, referred publicly to feedback as an undesired coupling between components of a vacuum-tube electronic system. The occasion was the lecture Braun delivered as a recipient of the Nobel Prize in Physics, shared with Guglielmo Marconi (often solely credited as the inventor of radio).

In 1927, Harold Black, at Bell Labs, invented the negative-feedback amplifier, which he described in detail in a seminal paper, "Stabilized Feedback Amplifiers," published in 1934. This invention was motivated by the need to provide low-distortion amplifiers that could be concatenated in long-distance transcontinental telephone circuits.

11.3.3 Analysis of the Feedback Voltage Amplifier Utilizing the Loop Gain

The feedback analysis method studied in Section 11.1 *cannot* be directly applied to a practical feedback voltage amplifier such as those in Fig. 11.8. This is because the analysis method of Section 11.1 is predicated on the assumption that the feedback network does not load the basic amplifier. Unfortunately, this assumption does not hold in most practical amplifier circuits. As shown in the circuits of Fig. 11.8, the feedback network is a simple resistive circuit that obviously loads the basic amplifier. As an example, in the circuit of Fig. 11.8(b), the values of the resistances R_2 and R_1, which comprise the feedback network, affect the gain of the common-source stage Q_2, which is part of the basic amplifier. Also, the value of the feedback-network resistance R_1 affects the gain of the Q_1 amplifier stage, which is part of the basic amplifier. It follows that we cannot easily disassemble a practical amplifier circuit to determine A and β and thus be able to use the feedback formulas of Sections 11.1 and 11.2.

While it is not easy to determine A and β, their product, the loop gain $A\beta$, can always be determined using the method presented in Section 11.1.3. Also, we can easily obtain the value of β by identifying and isolating the feedback network (e.g., the resistive divider (R_1, R_2) in each of the circuits in Fig. 11.8). We can then use the values of $A\beta$ and β to determine A and A_f. This loop-gain method is simple, and we shall use it in this section to perform the analysis of the feedback voltage amplifier. The method, however, has limitations that will be mentioned later. A more accurate and systematic approach for the analysis of feedback voltage amplifiers will be presented in the next section.

The loop-gain analysis method comprises four steps:

1. Identify the feedback network and use it to determine the value of β.
2. Determine the ideal value of the closed-loop gain A_f as $1/\beta$. This value of A_f is approached when $A\beta \gg 1$. The ideal or upper-bound value of A_f can be used in the initial design of the feedback amplifier. It also serves as a check on the actual value of A_f calculated below.
3. Use the method described in Section 11.1.3 to determine the loop gain $A\beta$. Recall that in breaking the loop, care should be taken to not change the conditions in the loop. Thus, if we break a feedback loop at XX′, as shown in Fig. 11.9(a), and apply a test voltage V_t to the terminals thus created to the left of XX′, the terminals to the right of XX′ must be connected to an impedance Z_t. The value of Z_t is equal

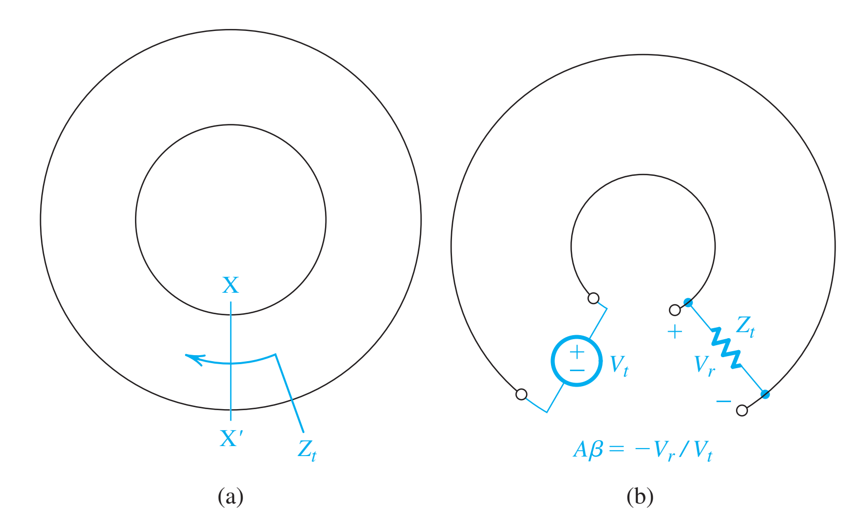


Figure 11.9 Breaking the conceptual feedback loop in (**a**) to determine the loop gain requires the termination of the loop as shown in (**b**), to ensure that the loop conditions do not change.

to the impedance previously seen looking to the left of XX′. The loop gain is then determined from

$$A\beta = -\frac{V_r}{V_t}$$

Whenever possible, we should break the loop at a location where Z_t is infinite.

4. Use the value of $A\beta$ together with that of β to determine the open-loop gain A. Then, determine the closed-loop gain A_f from

$$A_f = \frac{A}{1+A\beta}$$

We shall illustrate the application of the loop-gain method via two examples.

Example 11.2

For the series–shunt feedback amplifier of Fig. 11.8(b), neglect the MOSFETs' r_o and

(a) give the feedback network (β circuit) and an expression for β. Also give an expression for the ideal or upper-bound value of the closed-loop gain A_f.
(b) find the ratio R_2/R_1 that results in an ideal closed-loop gain of 10 V/V. If $R_1 = 1$ kΩ, what value must R_2 have?
(c) find an expression for the loop gain $A\beta$.
(d) if $g_{m1} = g_{m2} = 4$ mA/V and $R_{D1} = R_{D2} = 10$ kΩ, determine the values of $A\beta$, A, and A_f.

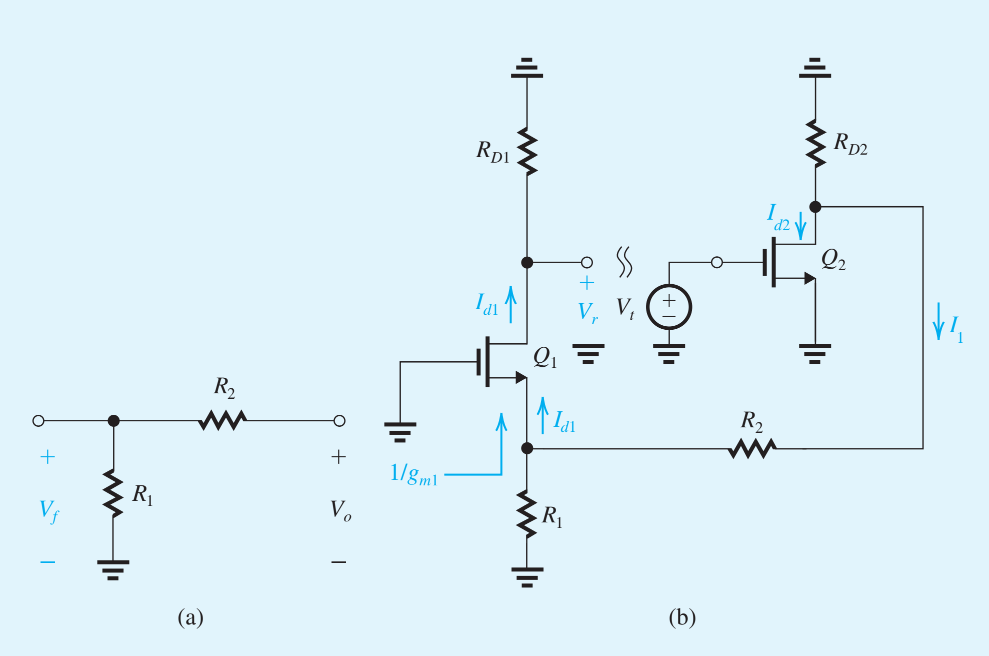


Figure 11.10 Determining: **(a)** the feedback factor β; and **(b)** the loop gain $A\beta$ for the feedback voltage amplifier of Fig. 11.8(b).

Solution

(a) The feedback network is highlighted in Fig. 11.8(b) and is redrawn in Fig. 11.10(a). It is a simple resistive voltage divider. Thus,

$$\beta = \frac{V_f}{V_o} = \frac{R_1}{R_1 + R_2}$$

For $A\beta \gg 1$, the closed-loop gain $A_f = V_o/V_s$ is given by

$$A_f \simeq \frac{1}{\beta} = 1 + \frac{R_2}{R_1}$$

This is the ideal or upper-bound value of A_f.

(b) For A_f to have an ideal value of 10,

$$10 = 1 + \frac{R_2}{R_1}$$

Thus,

$$\frac{R_2}{R_1} = 9$$

For $R_1 = 1\ \text{k}\Omega$, $R_2 = 9\ \text{k}\Omega$.

(c) To determine the loop gain, we set $V_s = 0$. Examining the feedback loop reveals that it is most convenient to break the loop at the connection between the drain of Q_1 and the gate of Q_2. This is

Example 11.2 *continued*

because of the infinite input impedance at the gate of Q_2. The resulting circuit is shown in Fig. 11.10(b), for which the analysis to determine $A\beta \equiv -V_r/V_t$ proceeds as follows:

$$I_{d2} = g_{m2} V_t$$

$$I_1 = -I_{d2} \frac{R_{D2}}{R_{D2} + R_2 + \left(R_1 \middle\| \frac{1}{g_{m1}}\right)}$$

$$I_{d1} = I_1 \frac{R_1}{R_1 + \frac{1}{g_{m1}}}$$

$$V_r = I_{d1} R_{D1}$$

Combining these four equations results in

$$A\beta \equiv -\frac{V_r}{V_t} = (g_{m1}R_{D1})(g_{m2}R_{D2}) \frac{1}{1 + g_{m1}R_1} \frac{R_1}{R_{D2} + R_2 + \left(R_1 \middle\| \frac{1}{g_{m1}}\right)}$$

(d)

$$A\beta = 4 \times 10 \times 4 \times 10 \times \frac{1}{1 + 4 \times 1} \times \frac{1}{10 + 9 + (1 \| \frac{1}{4})}$$

$$= 16.67$$

Since

$$\beta = \frac{R_1}{R_1 + R_2} = \frac{1}{1 + 9} = 0.1$$

$$A = \frac{A\beta}{\beta} = \frac{16.67}{0.1} = 166.7 \text{ V/V}$$

and

$$A_f = \frac{A}{1 + A\beta} = \frac{166.7}{1 + 16.67} = \frac{166.7}{17.67} = 9.43 \text{ V/V}$$

which, given that the loop gain (16.67) is only moderately high, is reasonably close to the ideal value of 10 V/V.

Example 11.3

In the series–shunt feedback amplifier of Fig. 11.11(a), the op amp has an input resistance R_{id}, an open-circuit voltage gain μ, and an output resistance r_o. Find expressions for β, the ideal value of $A_f \equiv V_o/V_s$, and the loop gain $A\beta$.

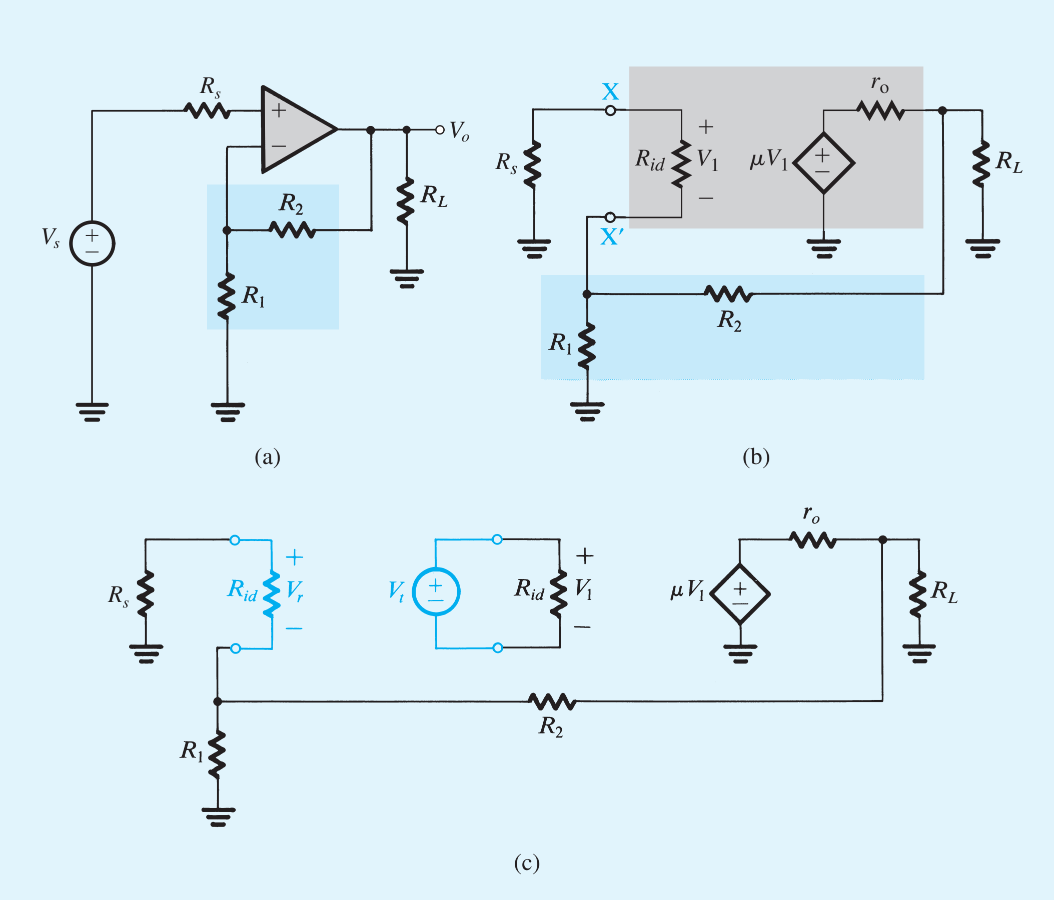


Figure 11.11 Example 11.3. (**a**) A series–shunt feedback amplifier; (**b**) the feedback loop obtained by setting $V_s = 0$ and replacing the op amp with its equivalent-circuit model; (**c**) breaking the feedback loop to determine the loop gain $A\beta = -V_r/V_t$.

Solution

The feedback network consists of the voltage divider (R_1, R_2), thus

$$\beta = \frac{R_1}{R_1 + R_2}$$

and the ideal value of A_f is

$$A_f = \frac{1}{\beta} = 1 + \frac{R_2}{R_1}$$

To determine the loop gain, we set $V_s = 0$ and replace the op amp with its equivalent-circuit model. The resulting circuit is shown in Fig. 11.11(b). Next, we break the loop to apply a test voltage V_t while terminating the loop at the break in an impedance equal to that seen prior to breaking the loop. The resulting circuit is shown in Fig. 11.11(c), where the loop has been broken at the input terminals of the op amp and the left-hand-side terminals are connected to a resistance equal to R_{id}. Analysis of the circuit to determine

Example 11.3 *continued*

$A\beta \equiv -V_r/V_t$ involves repeated application of the voltage divider rule, resulting in

$$A\beta = \mu \frac{\{R_L \parallel [R_2 + R_1 \parallel (R_{id} + R_s)]\}}{\{R_L \parallel [R_2 + R_1 \parallel (R_{id} + R_s)]\} + r_o} \times \frac{[R_1 \parallel (R_{id} + R_s)]}{[R_1 \parallel (R_{id} + R_s)] + R_2} \times \frac{R_{id}}{R_{id} + R_s}$$

EXERCISE

11.6 For the feedback voltage amplifier of Fig. 11.8(c):

(a) Find an expression for β.

(b) Neglecting the MOSFET r_o, find an expression for the loop gain $A\beta$. (*Hint*: Break the loop at the gate of Q.)

(c) Find an expression for the open-loop gain A.

(d) For $g_m = 4$ mA/V, $R_D = 10$ kΩ, $R_1 = 20$ kΩ, and $R_2 = 80$ kΩ, find the values of β, $A\beta$, A, and the closed-loop gain A_f. What would A_f be if $A\beta$ were much greater than unity?

Ans. (a) $\frac{R_1}{R_1 + R_2}$; (b) $g_m \frac{R_D R_1}{R_D + R_1 + R_2}$; (c) $g_m \frac{R_D(R_1 + R_2)}{R_D + R_1 + R_2}$; (d) 0.2, 7.27, 36.36 V/V, 4.4 V/V, 5 V/V.

11.3.4 A Final Remark

The loop-gain analysis method, though simple, is not complete: It does not enable us to find the input and output resistances of the feedback amplifier. This shortcoming is remedied in the next section, where we present a systematic approach to the analysis of feedback voltage amplifiers.

11.4 Systematic Analysis of Feedback Voltage Amplifiers

In this section we provide a systematic procedure for the analysis of feedback voltage amplifiers. The procedure essentially disassembles a given feedback voltage-amplifier circuit so as to obtain the "A circuit," from which we can determine the open-loop gain A and other parameters of the open-loop amplifier, such as the input and output resistances, and the "β circuit" from which the value of the feedback factor β can be found. We can then use the

feedback formulas to determine the characteristic parameters of the feedback amplifier, such as the closed-loop gain A_f and the input and output resistances with feedback.

Our approach will be to first consider the *ideal* case in which the feedback network does *not* load the basic amplifier. Then, we consider the practical case in which not only does the feedback network load the basic amplifier, but also there is a finite source resistance R_s and a finite load resistance R_L.

11.4.1 The Ideal Case

As mentioned before, series–shunt is the appropriate feedback topology for a voltage amplifier. The ideal structure of the series–shunt feedback amplifier is shown in Fig. 11.12(a). It consists of a *unilateral* open-loop amplifier (the A circuit) and an ideal voltage-sampling, voltage-mixing feedback network (the β circuit). The A circuit has an input resistance R_i, an open-circuit voltage gain A, and an output resistance R_o. It is assumed that the source is ideal with a zero resistance and that there is no load resistance. Furthermore, note that the β circuit

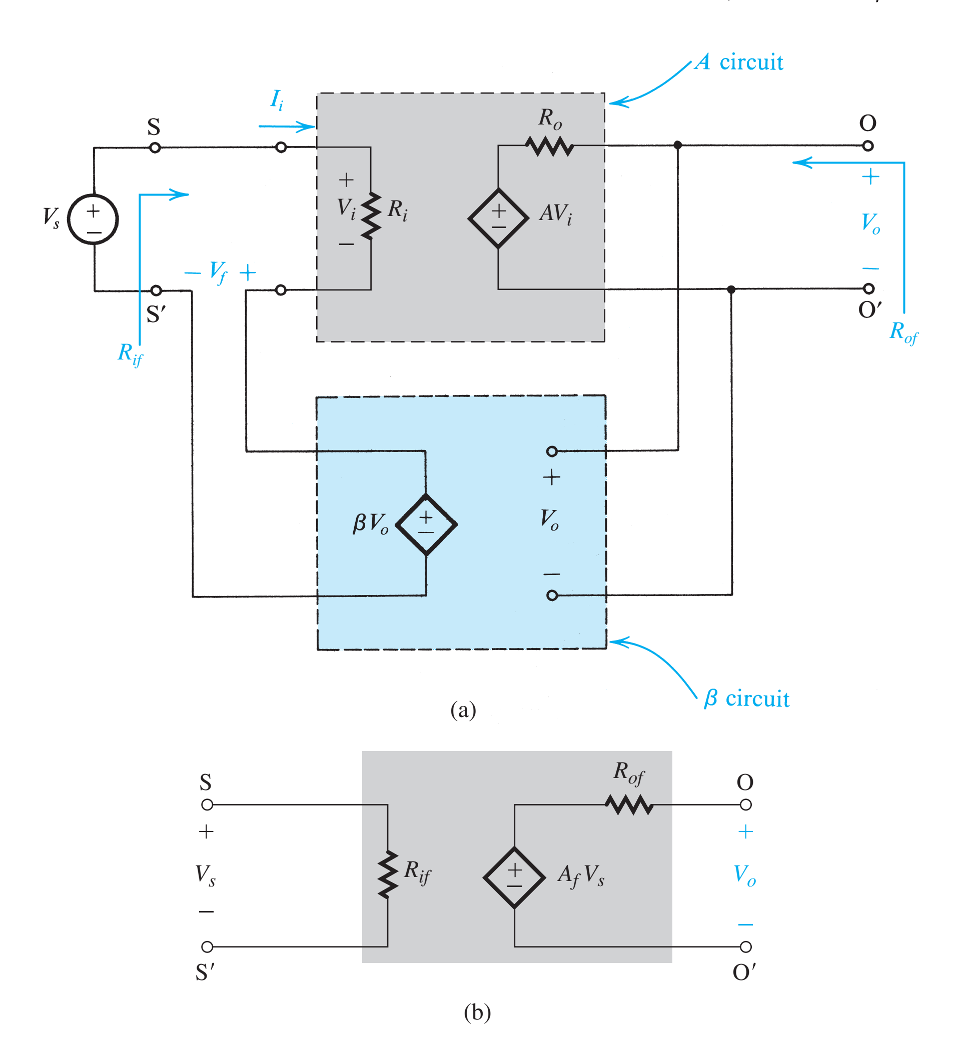


Figure 11.12 The series–shunt feedback amplifier: **(a)** ideal structure; **(b)** equivalent circuit.

does *not* load the A circuit; that is, connecting the β circuit does not change the value of A (defined as $A \equiv V_o/V_i$).

The circuit of Fig. 11.12(a) exactly follows the ideal feedback model of Fig. 11.1. Therefore the closed-loop voltage gain A_f is given by

$$A_f \equiv \frac{V_o}{V_s} = \frac{A}{1+A\beta} \tag{11.18}$$

The equivalent-circuit model of the series–shunt feedback amplifier is shown in Fig. 11.12(b). Observe that A_f is the open-circuit voltage gain of the feedback amplifier, R_{if} is its input resistance, and R_{of} is its output resistance. Expressions for R_{if} and R_{of} can be derived as follows.

For R_{if}, refer to the input loop of the circuit in Fig. 11.12(a). The series mixing subtracts V_f from V_s and thus reduces V_i by a factor equal to the amount of feedback (Eq. 11.7),

$$V_i = \frac{V_s}{1+A\beta}$$

Thus the input current I_i becomes

$$I_i = \frac{V_i}{R_i} = \frac{V_s}{(1+A\beta)R_i} \tag{11.19}$$

Since I_i is the current drawn from V_s, the input resistance R_{if} can be expressed as

$$R_{if} \equiv \frac{V_s}{I_i}$$

and using Eq. (11.19) is found to be

$$R_{if} = (1+A\beta)R_i \tag{11.20}$$

Thus, as expected, the series-mixing feedback results in an increase in the amplifier input resistance by a factor equal to the amount of feedback, $(1+A\beta)$, a highly desirable property for a voltage amplifier.

To determine the output resistance R_{of} of the feedback amplifier in Fig. 11.12(a), we set $V_s = 0$ and apply a test voltage V_x between the output terminals, as shown in Fig. 11.13. If the current drawn from V_x is I_x, the output resistance R_{of} is

$$R_{of} \equiv \frac{V_x}{I_x} \tag{11.21}$$

An equation for the output loop yields

$$I_x = \frac{V_x - AV_i}{R_o} \tag{11.22}$$

From the input loop we see that

$$V_i = -V_f$$

Now $V_f = \beta V_o = \beta V_x$; thus,

$$V_i = -\beta V_x$$

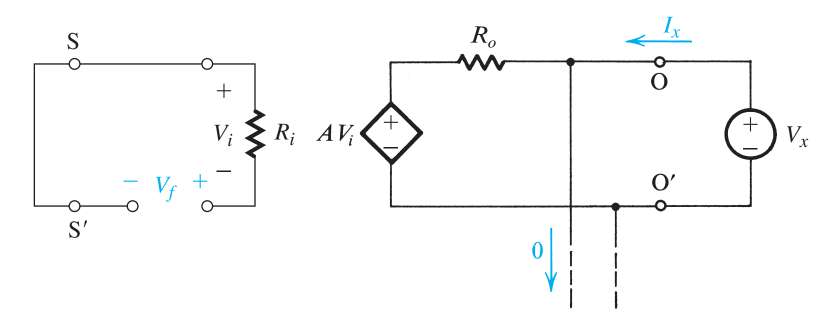


Figure 11.13 Determining the output resistance of the feedback amplifier of Fig. 11.12(a): $R_{of} = V_x/I_x$.

which when substituted in Eq. (11.22) yields

$$I_x = \frac{V_x(1+A\beta)}{R_o}$$

Substituting this value of I_x into Eq. (11.21) provides the following expression for R_{of},

$$R_{of} = \frac{R_o}{1+A\beta} \tag{11.23}$$

Thus, as expected, the shunt sampling (or voltage sampling) at the output results in a decrease in the amplifier output resistance by a factor equal to the amount of negative feedback, $(1+A\beta)$, another highly desirable property for a voltage amplifier.

Although perhaps not entirely obvious, the reduction of the output resistance is a result only of the method of sampling the output and does not depend on the method of mixing. Thus, the transistance amplifier, which is the other amplifier type in which shunt (or voltage) sampling is employed, will also exhibit a reduced output resistance.

11.4.2 The Practical Case

In a practical series–shunt feedback amplifier, the feedback network will not be an ideal voltage-controlled voltage source. Rather, as in the circuits of Fig. 11.8, the feedback network is usually resistive and hence will load the basic amplifier and thus affect the values of A, R_i, and R_o. In addition, there will be finite source and load resistances, which in turn will affect these three parameters. Thus the problem we have is as follows: Given a series–shunt feedback amplifier represented by the block diagram of Fig. 11.14(a), find the A circuit and the β circuit.

The problem in essence is to represent the general feedback voltage amplifier of Fig. 11.14(a) with the ideal structure of Fig. 11.12(a). The solution is presented, without derivation, in Fig. 11.14(b). We make the following observations.

1. The A circuit is obtained by augmenting the basic amplifier at the input with the source resistance R_s and a resistance R_{11}, and at the output with the load resistance R_L and a resistance R_{22}. Resistances R_{11} and R_{22} represent the loading effect of the feedback network on the basic amplifier at the input and the output, respectively.
2. Resistance R_{11} is the resistance looking into port 1 of the feedback network while port 2 is short-circuited. Resistance R_{22} is the resistance looking into port 2 of the

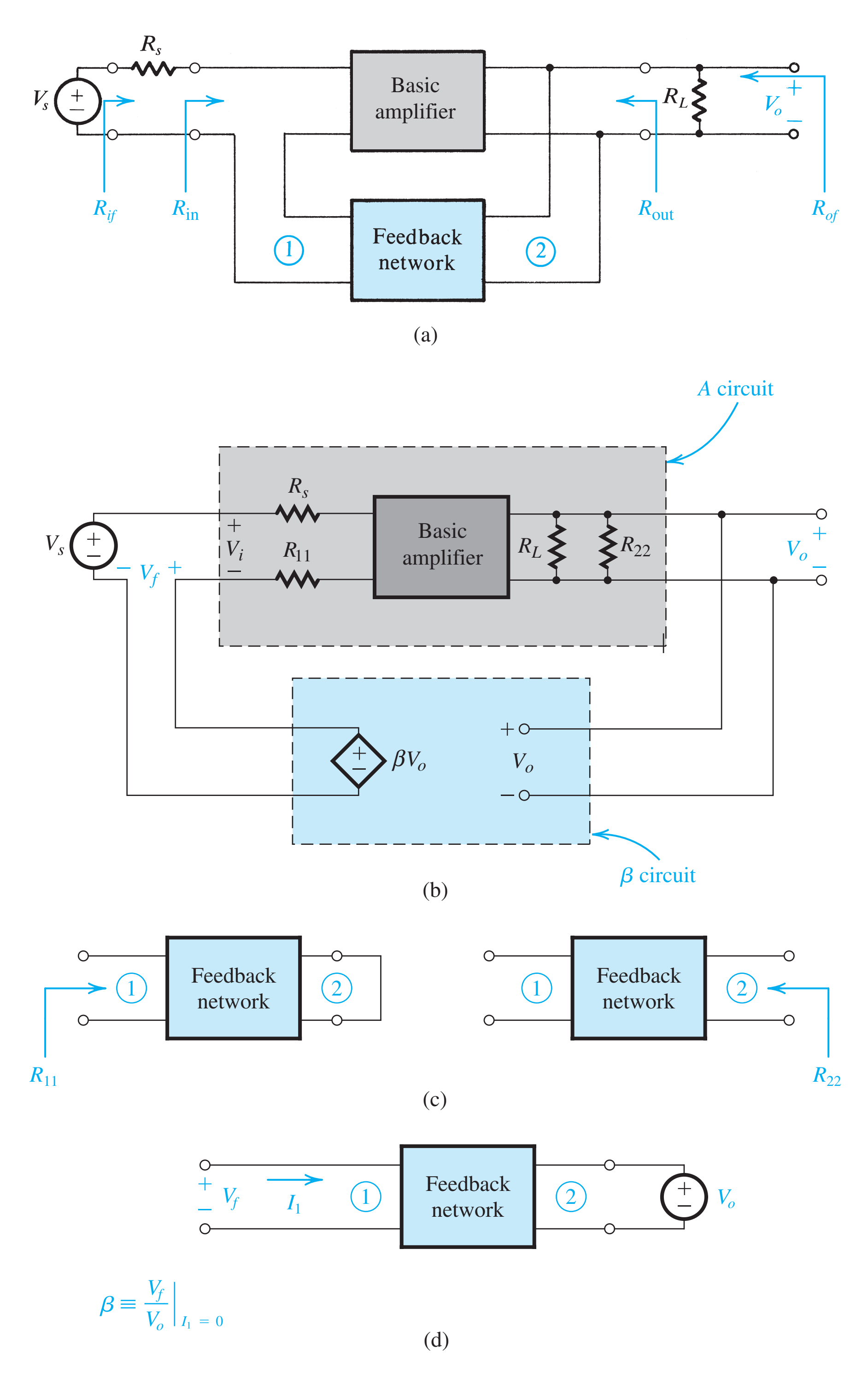


Figure 11.14 **(a)** Block diagram of a practical series–shunt feedback amplifier. **(b)** The circuit in **(a)** represented by the ideal structure of Fig. 11.12(a). **(c)** Definition of R_{11} and R_{22}. **(d)** Determination of the feedback factor β. **(e)** The A circuit, showing the open-loop resistances R_i and R_o.

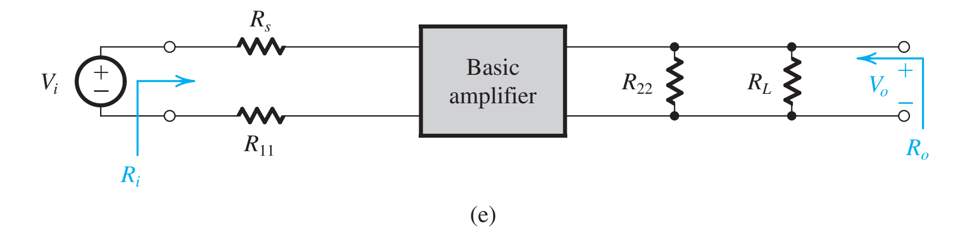


Figure 11.14 *continued*

feedback network while port 1 is open-circuited. These definitions are illustrated in Fig. 11.14(c). Since the feedback network is connected in shunt with the output, shorting its port 2 destroys the feedback. Similarly, because the feedback network is connected in series with the input, opening its port 1 destroys the feedback. It follows that the loading effect of the feedback network is obtained by looking into its appropriate port while the other port is open-circuited or short-circuited so as to destroy the feedback.[4]

3. The feedback factor β is the transmission from port 2 to port 1 of the feedback network, with port 1 open-circuited (which destroys the feedback). Reference to Fig. 11.14(c) shows that

$$\beta \equiv \left.\frac{V_f}{V_o}\right|_{I_1=0} \tag{11.24}$$

4. The open-loop gain A can be obtained from the A circuit in Fig. 11.14(e) as $A = \dfrac{V_o}{V_i}$

5. The values of A and β can be used to determine the closed-loop gain A_f,

$$A_f \equiv \frac{V_o}{V_s} = \frac{A}{1 + A\beta}$$

6. The open-loop input resistance R_i and output resistance R_o can be determined from the A circuit [see Fig. 11.14(e)]. These values can be used to determine the input and output resistances with feedback,

$$R_{if} = R_i(1 + A\beta)$$
$$R_{of} = R_o/(1 + A\beta)$$

From Fig. 11.14(a) we see that R_{if} is the resistance seen by the ideal signal source V_s. The actual input resistance of the feedback amplifier R_{in} excludes R_s and is found from R_{if},

$$R_{\text{in}} = R_{if} - R_s \tag{11.25}$$

[4]A simple rule to remember: If the connection is *sh*unt, *sh*ort it; if *se*ries, *se*ver it.

Similarly, R_{of} is the output resistance of the feedback amplifier including R_L. The actual output resistance excludes R_L (see Fig. 11.14(a)) and is found from R_{of},

$$R_{\text{out}} = 1\Big/\left(\frac{1}{R_{of}} - \frac{1}{R_L}\right) \tag{11.26}$$

A final and important note: The representation in Fig. 11.14(b) is only *approximately* equivalent to the original circuit in Fig. 11.14(a). The approximation is a result of neglecting the small forward transmission in the feedback network relative to the much larger forward transmission in the basic amplifier. Also, recall that we continue to assume that the basic amplifier is unilateral—that is, it does not have internal feedback; all the feedback occurs in the feedback network and is represented by the feedback factor β.

Example 11.4

Figure 11.15(a) shows an op amp connected in the noninverting configuration. The op amp has an open-circuit voltage gain μ, a differential input resistance R_{id}, and an output resistance r_o. Recall that in our analysis of op-amp circuits in Chapter 2, we neglected the effects of R_{id} (assumed it to be infinite) and of r_o (assumed it to be zero). Here we wish to use the feedback method to analyze the circuit taking both R_{id} and r_o into account. Find expressions for A, β, the closed-loop gain V_o/V_s, the input resistance R_{in} [see Fig. 11.15(a)], and the output resistance R_{out}. Also find numerical values, given $\mu = 10^4$, $R_{id} = 100\ \text{k}\Omega$, $r_o = 1\ \text{k}\Omega$, $R_L = 2\ \text{k}\Omega$, $R_1 = 1\ \text{k}\Omega$, $R_2 = 1\ \text{M}\Omega$, and $R_s = 10\ \text{k}\Omega$. Note that this circuit was analyzed in Example 11.3 using the loop-gain method and, where appropriate, compare results.

Solution

We observe that the feedback network consists of R_2 and R_1. This network samples the output voltage V_o and provides a voltage signal (across R_1) that is mixed in series with the input source V_s.

The A circuit can be easily obtained following the rules of Fig. 11.14, and is shown in Fig. 11.15(b). Observe that the loading effect of the feedback network at the input side is obtained by short-circuiting port 2 of the feedback network (because it is connected in shunt) and looking into port 1, with the result that $R_{11} = R_1 \| R_2$. The loading effect of the feedback network at the output side is found by open-circuiting port 1 (because it is connected in series) and looking into port 2, with the result that $R_{22} = R_2 + R_1$. For the resulting A circuit in Fig. 11.15(b), we can write by inspection:

$$A \equiv \frac{V_o}{V_i} = \mu \frac{R_L \| (R_1 + R_2)}{[R_L \| (R_1 + R_2)] + r_o} \frac{R_{id}}{R_{id} + R_s + (R_1 \| R_2)}$$

For the values given, we find that $A \simeq 6000$ V/V.

The circuit for determining β is shown in Fig. 11.15(c), from which we obtain

$$\beta \equiv \frac{V_f}{V_o} = \frac{R_1}{R_1 + R_2} \simeq 10^{-3}\ \text{V/V}$$

Thus,

$$A\beta = 6000 \times 10^{-3} = 6$$

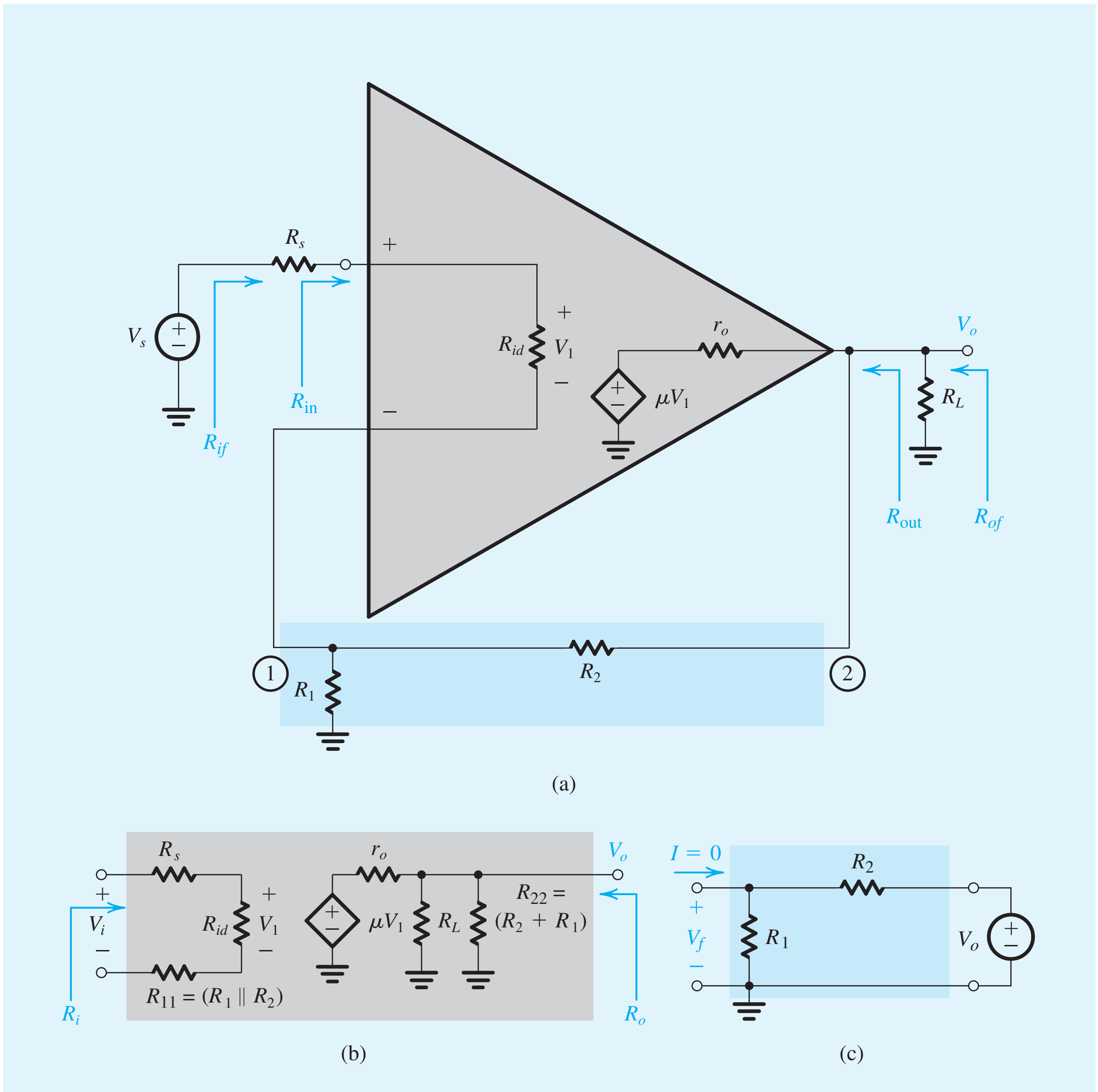


Figure 11.15 Circuits for Example 11.4.

and

$$1 + A\beta = 7$$

Identical results are obtained by substituting the given numerical values into the expression for $A\beta$ derived in Example 11.3.

The voltage gain with feedback can now be obtained as

$$A_f \equiv \frac{V_o}{V_s} = \frac{A}{1 + A\beta} = \frac{6000}{7} = 857 \text{ V/V}$$

Example 11.4 *continued*

The input resistance R_{if} determined by the feedback equations is the resistance seen by the external source (see Fig. 11.15a), and is given by

$$R_{if} = R_i(1 + A\beta)$$

where R_i is the input resistance of the A circuit in Fig. 11.15(b):

$$R_i = R_s + R_{id} + (R_1 \parallel R_2)$$

For the values given, $R_i \simeq 111$ kΩ, resulting in

$$R_{if} = 111 \times 7 = 777 \text{ k}\Omega$$

This, however, is not the resistance asked for. What is required is R_{in}, indicated in Fig. 11.15(a). To obtain R_{in} we subtract R_s from R_{if}:

$$R_{\text{in}} = R_{if} - R_s = 767 \text{ k}\Omega$$

The resistance R_{of} given by the feedback equations is the output resistance of the feedback amplifier, including the load resistance R_L, as indicated in Fig. 11.15(a). R_{of} is given by

$$R_{of} = \frac{R_o}{1 + A\beta}$$

where R_o is the output resistance of the A circuit. R_o can be obtained by inspection of Fig.11.15(b) with V_i set to zero, as

$$R_o = r_o \parallel R_L \parallel \left(R_2 + R_1\right)$$

For the values given, $R_o \simeq 666$ Ω, and

$$R_{of} = \frac{666}{7} = 95.2 \ \Omega$$

The resistance asked for, R_{out}, is the output resistance of the feedback amplifier excluding R_L. From Fig. 11.15(a) we see that

$$R_{of} = R_{\text{out}} \parallel R_L$$

Thus

$$R_{\text{out}} \simeq 100 \ \Omega$$

Example 11.5

As another example of a series–shunt feedback amplifier, consider the circuit shown in Fig. 11.8(b), which we analyzed in Example 11.2 by determining the loop gain $A\beta$. In this example we wish to first analyze the circuit using our systematic procedure and then compare the results to those obtained in Example 11.2. For convenience, the circuit is repeated in Fig. 11.16(a). It is required to obtain the voltage gain V_o/V_s, input resistance R_{in}, and output resistance R_{out}. Find numerical values for the case $g_{m1} = g_{m2} = 4$ mA/V, $R_{D1} = R_{D2} = 10$ kΩ, $R_1 = 1$ kΩ, and $R_2 = 9$ kΩ. For simplicity, neglect r_o of each of Q_1 and Q_2.

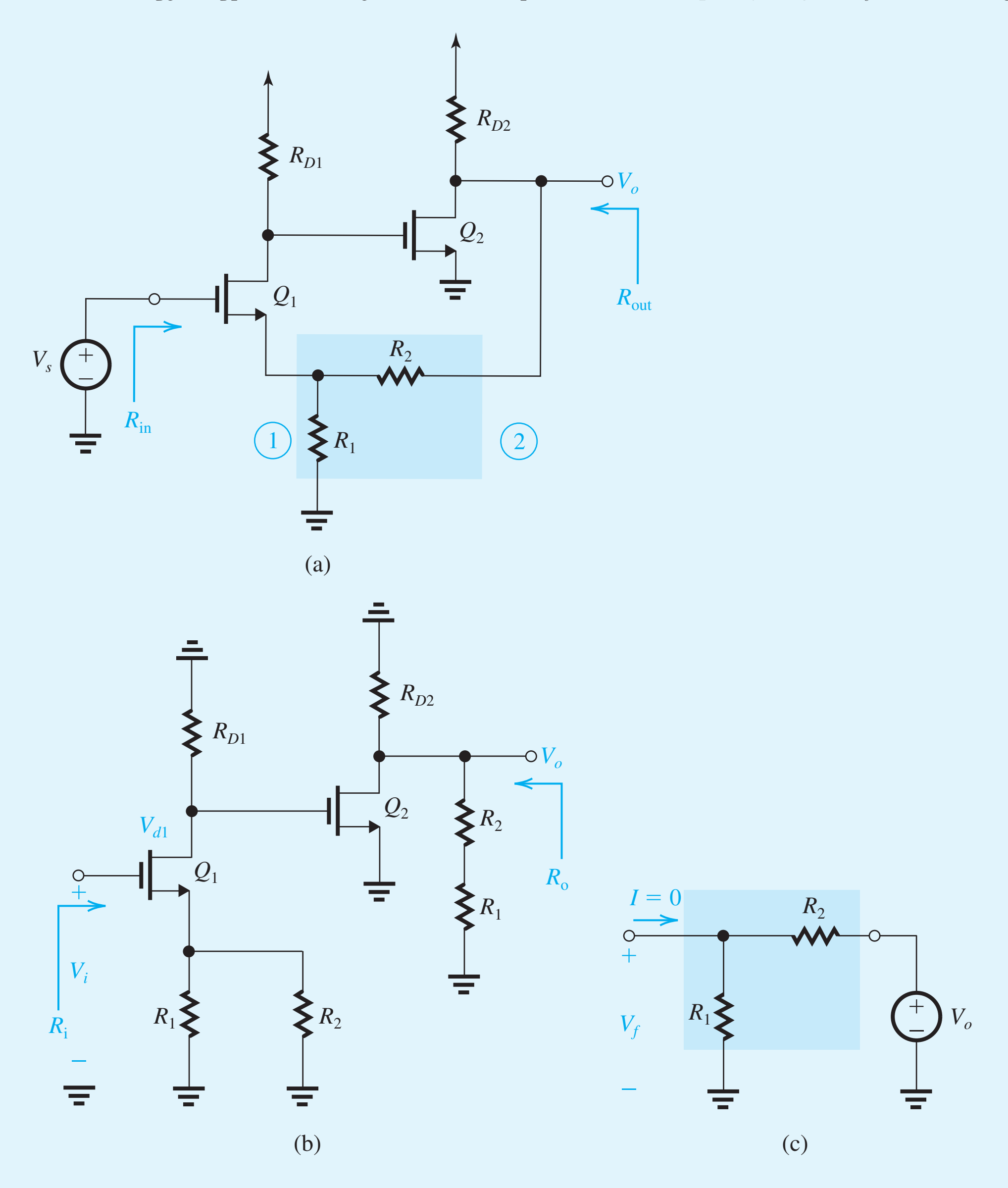


Figure 11.16 **(a)** Series–shunt feedback amplifier for Example 11.5; **(b)** the *A* circuit; **(c)** the β circuit.

Example 11.5 *continued*

Solution

We identify the feedback network as the voltage divider (R_1, R_2). Its loading effect at the input is obtained by short-circuiting its port 2 (because it is connected in shunt with the output). Then, looking into its port 1, we see $R_1 \| R_2$. The loading effect at the output is obtained by open-circuiting port 1 of the feedback network (because it is connected in series with the input). Then, looking into port 2, we see R_2 in series with R_1. The A circuit will therefore be as shown in Fig. 11.16(b). The gain A is determined as the product of the gain of Q_1 and the gain of Q_2 as follows:

$$A_1 = \frac{V_{d1}}{V_i} = -\frac{R_{D1}}{1/g_{m1} + (R_1 \| R_2)} = -\frac{g_{m1}R_{D1}}{1 + g_{m1}(R_1 \| R_2)}$$

$$A_2 = \frac{V_o}{V_{d1}} = -g_{m2}[R_{D2} \| (R_1 + R_2)]$$

$$A = \frac{V_o}{V_i} = A_1 A_2 = \frac{g_{m1}R_{D1}g_{m2}[R_{D2} \| (R_1 + R_2)]}{1 + g_{m1}(R_1 \| R_2)}$$

For the numerical values given,

$$A = \frac{4 \times 10 \times 4[10 \| (1+9)]}{1 + 4(1 \| 9)} = 173.9 \text{ V/V}$$

which is reasonably close to the value of 166.7 V/V obtained in Example 11.2.

The value of β is determined from the β circuit in Fig. 11.16(c),

$$\beta \equiv \frac{V_f}{V_o} = \frac{R_1}{R_1 + R_2}$$

For the numerical values given,

$$\beta = \frac{1}{1+9} = 0.1$$

The closed-loop gain V_o/V_s can now be found as

$$\frac{V_o}{V_s} = A_f = \frac{A}{1 + A\beta} = \frac{173.9}{1 + 173.9 \times 0.1} = 9.46 \text{ V/V}$$

which is very close to the value of 9.43 V/V obtained in Example 11.2.

The input resistance is obviously infinite because of the infinite input resistance of the MOSFET. The output resistance R_{out} is obtained as follows,

$$R_{out} = R_{of} = \frac{R_o}{1 + A\beta}$$

where R_o is the output resistance of the A circuit. From Fig. 11.16(b),

$$R_o = R_{D2} \| (R_1 + R_2)$$
$$= 10 \| 10 = 5\,\text{k}\Omega$$

The amount of feedback is

$$1 + A\beta = 1 + (173.9 \times 0.1) = 18.39$$

Thus,

$$R_{\text{out}} = \frac{5000}{18.39} = 272\ \Omega$$

which is relatively low given that the open-loop amplifier has $R_o = 5000\ \Omega$. We finally note that the loop-gain method utilized in Example 11.2 cannot provide the input and output resistances because the A circuit is not determined.

EXERCISES

11.7 If the op amp of Example 11.4 has a uniform –6-dB/octave high-frequency rolloff with $f_{3\text{dB}} = 1$ kHz, find the 3-dB frequency of the closed-loop gain V_o/V_s.
Ans. 7 kHz

11.8 The circuit shown in Fig. E11.18 consists of a differential stage followed by an emitter follower, with series–shunt feedback supplied by the resistors R_1 and R_2. Assuming that the dc component of V_s is zero, and that β of the BJTs is very high, find the dc operating current of each of the three transistors and show that the dc voltage at the output is approximately zero. Then find the values of A, β, $A_f \equiv V_o/V_s$, R_{in}, and R_{out}. Assume that the transistors have $\beta = 100$.
Ans. 0.5 mA, 0.5 mA, 5 mA; 85.7 V/V; 0.1 V/V; 8.96 V/V; 191 kΩ; 19.1 Ω.

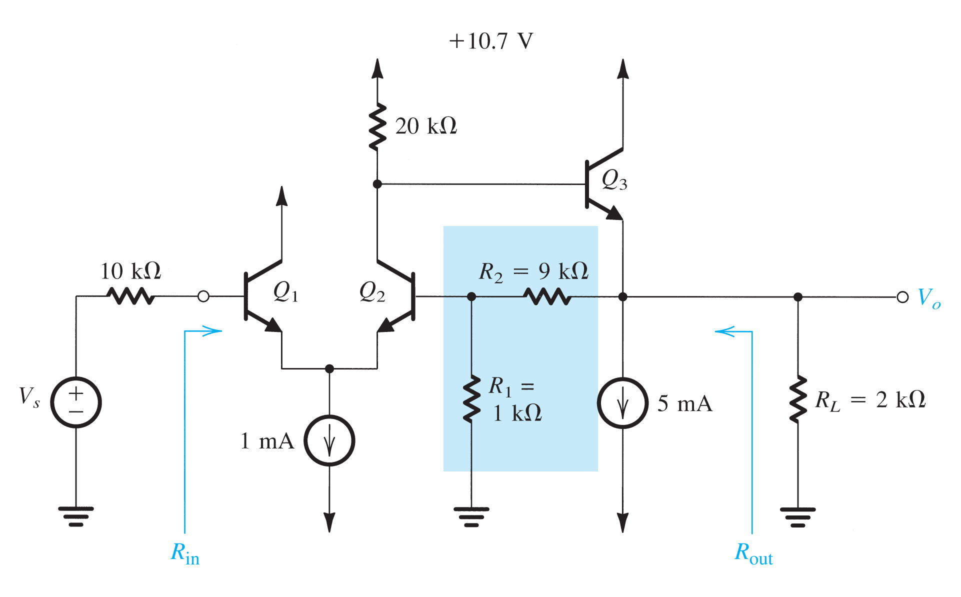


Figure E11.8

11.9 For the series–shunt amplifier in Fig. 11.8(c), which was considered in Exercise 11.6, find A, β, A_f, R_{in}, and R_{out}. Neglect r_o of Q. Compare results to those obtained in Exercise 11.6.

Ans. $A = g_m\left[R_D \parallel (R_1 + R_2)\right]$; $\beta = R_1/(R_1 + R_2)$;
$A_f = A/(1 + A\beta)$; $R_{in} = (1/g_m)(1 + A\beta)$;
$R_{out} = \left[R_D \parallel (R_1 + R_2)\right]/(1 + A\beta)$

Comparison: A and β are identical to the corresponding expressions found in Exercise 11.6. However, R_{in} and R_{out} cannot be determined using the method of Exercise 11.6.

11.5 Other Feedback-Amplifier Types

Having studied in detail the most common feedback-amplifier type, the feedback voltage amplifier, we now consider the three other types of feedback amplifier: the feedback transconductance amplifier, the feedback current amplifier, and the feedback transresistance amplifier. The presentation builds on that of the feedback voltage amplifier, and the results will be given without derivation. The analysis method will be illustrated with a large number of worked-out examples dealing with practical and widely utilized circuits.

11.5.1 Basic Principles

1. *Sensing:* The feedback network must sample the output signal of interest. Thus if V_o is the output signal of interest, as in the case of voltage and transresistance amplifiers, the feedback network is connected in parallel (or *shunt*) with the amplifier output node, just as a voltmeter is connected to measure a voltage. On the other hand, if I_o is the output signal of interest, as in the case of transconductance and current amplifiers, the feedback network is connected in *series* with the output loop, just as a current meter is connected to measure a current.
2. *Mixing:* If the input signal to be amplified is a voltage, as in the case of voltage and transconductance amplifiers, the signal source is represented by its Thévenin equivalent and the feedback voltage signal V_f is connected in *series* with the input signal source V_s. On the other hand, if the signal to be amplified is a current, as in the case of current and transresistance amplifiers, the Norton form is used to represent the signal source, and the feedback current signal I_f is connected in parallel (*shunt*) with the input signal source I_s.
3. *Feedback topology:* From the above, it follows that for each of the four amplifier types there is a uniquely appropriate feedback topology:

Amplifier Type	Appropriate Feedback Topology
Voltage	Series–Shunt
Transconductance	Series–Series
Current	Shunt–Series
Transresistance	Shunt–Shunt

The appropriate feedback topology not only stabilizes the gain of interest (e.g., the transconductance $A_f \equiv I_o/V_s$ in a transconductance amplifier), but also makes the input and output resistances more ideal (e.g., the shunt–series topology decreases the input resistance and increases the output resistance of a current amplifier).

4. *Input and output resistance:* The increase or decrease of the input or output resistance depends *solely* on the type of connection: Series connection always increases the resistance; parallel (shunt) connection always decreases the resistance. Furthermore, the increase or decrease is always by the amount of feedback, $(1 + A\beta)$. Thus, as an example, for the feedback current amplifier, the shunt connection at the input decreases the input resistance; $R_{if} = R_i/(1 + A\beta)$, and the series connection at the output increases the output resistance; $R_{of} = (1 + A\beta)R_o$, where R_i and R_o are the input and output resistances of the open-loop amplifier (A circuit).
5. *Dimensions of $A, \beta, A\beta$, and A_f:* Depending on the amplifier type, A, β, and A_f have the dimensions of V/V, A/A, V/A, or A/V. However, $A\beta$ is always dimensionless. For a feedback transconductance amplifier, for example, $A \equiv I_o/V_i$ (A/V), $\beta \equiv V_f/I_o$ (V/A), $A_f \equiv I_o/V_s$ (A/V), and $A\beta$ is in V/V or essentially dimensionless.
6. *Analysis using the loop gain:* For any feedback-amplifier type, the loop gain $A\beta$ can be determined using the method described in Section 11.3.3. The loop gain can then be used together with the feedback factor β to determine the open-loop gain A and hence the closed-loop gain A_f. This approach, however, does not enable the determination of the input and output resistances of the feedback amplifier. For these, we need to obtain the A circuit using the systematic analysis approach described below.

Example 11.6

Figure 11.17(a) shows a feedback transconductance amplifier utilizing an op amp together with an NMOS transistor. The feedback network consists of a resistor R_F that senses the output current I_o (recall that the drain and source currents of the MOSFET are equal) and provides a feedback voltage that is subtracted from V_s by means of the differencing action of the op-amp input. Observe that the feedback topology is series–series, which is uniquely appropriate for transconductance amplifiers.

(a) Find β and hence the closed-loop gain $A_f \equiv I_o/V_s$ obtained when $A\beta \gg 1$.

(b) Replace the op amp with its equivalent-circuit model characterized by an open-circuit voltage gain μ, an input differential resistance R_{id}, and an output resistance r_{o1}. Also, replace the MOSFET with its hybrid-π model characterized by a transconductance g_m and an output resistance r_{o2}. With the complete equivalent circuit in hand, set $V_s = 0$ and break the loop to determine $A\beta$.

(c) Use $A\beta$ together with β to find A.

(d) For $\mu = 1000$, $R_{id} = 100\ \text{k}\Omega$, $g_m = 2$ mA/V, $r_{o2} = 20\ \text{k}\Omega$, and $R_F = 1\ \text{k}\Omega$, find the values of β, $A\beta$, A, and A_f. Compare A_f to the ideal value obtained when $A\beta \gg 1$.

Example 11.6 *continued*

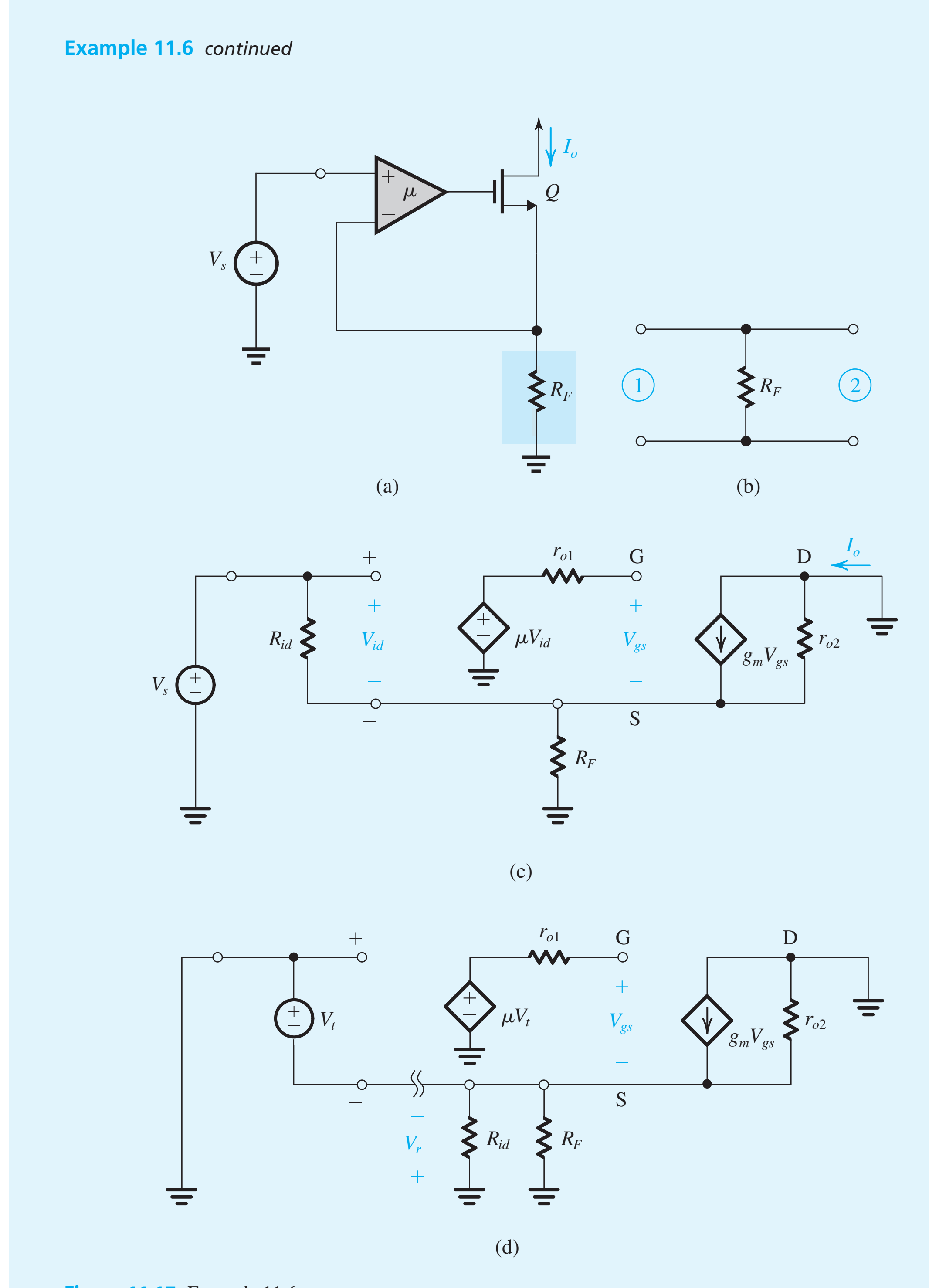


Figure 11.17 Example 11.6.

Solution

(a) The two-port feedback network is shown in Fig. 11.17(b). When port 2 is fed with a current I_o the voltage V_f appearing across port 1 is I_oR_F, thus

$$\beta = R_F$$

When $A\beta \gg 1$,

$$A_f \equiv \frac{I_o}{V_s} \simeq \frac{1}{\beta} = \frac{1}{R_F}$$

(b) Figure 11.7(c) shows the equivalent-circuit model of the feedback transconductance amplifier. To determine $A\beta$, we set $V_s = 0$ and break the loop at the input terminals of the op amp as shown in Fig. 11.17(d). Note that we have applied an input voltage V_t and terminated the loop at the break in the resistance R_{id} to prevent any change in the loop conditions. This circuit can be easily analyzed to determine $A\beta \equiv -V_r/V_t$ as follows:

$$V_{gs} = \mu V_t - (-V_r) = \mu V_t + V_r$$

$$-V_r = g_m V_{gs}(R_F \| R_{id} \| r_{o2})$$

Combining these two equations yields

$$A\beta \equiv -\frac{V_r}{V_t} = \mu \frac{g_m(R_F \| R_{id} \| r_{o2})}{1 + g_m(R_F \| R_{id} \| r_{o2})}$$

(c) Substituting $\beta = R_F$ provides

$$A = \frac{\mu}{R_F} \frac{g_m(R_F \| R_{id} \| r_{o2})}{1 + g_m(R_F \| R_{id} \| r_{o2})}$$

(d) For the given numerical values,

$$\beta = 1\ \text{k}\Omega$$

$$A\beta = 1000 \times \frac{2(1 \| 100 \| 20)}{1 + 2(1 \| 100 \| 20)}$$

$$= 653.6$$

$$A = 653.6\ \text{mA/V}$$

$$A_f \equiv \frac{I_o}{V_s} = \frac{653.6}{1 + 653.6} = 0.9985\ \text{mA/V}$$

Since the ideal value of A_f is 1 mA/V, the actual A_f is only 0.15% lower than ideal.

EXERCISES

11.10 For the circuit in Example 11.6, let the op-amp gain decrease by 10%. What is the corresponding percentage change in $A_f \equiv I_o/V_s$?
Ans. -0.02%

D11.11 Redesign the circuit in Example 11.6 to obtain a nominal closed-loop transconductance of 2 mA/V. What is the required value of R_F, and what is the actual transconductance A_f realized?
Ans. $R_F = 500\ \Omega$; $A_f = 1.996$ mA/V

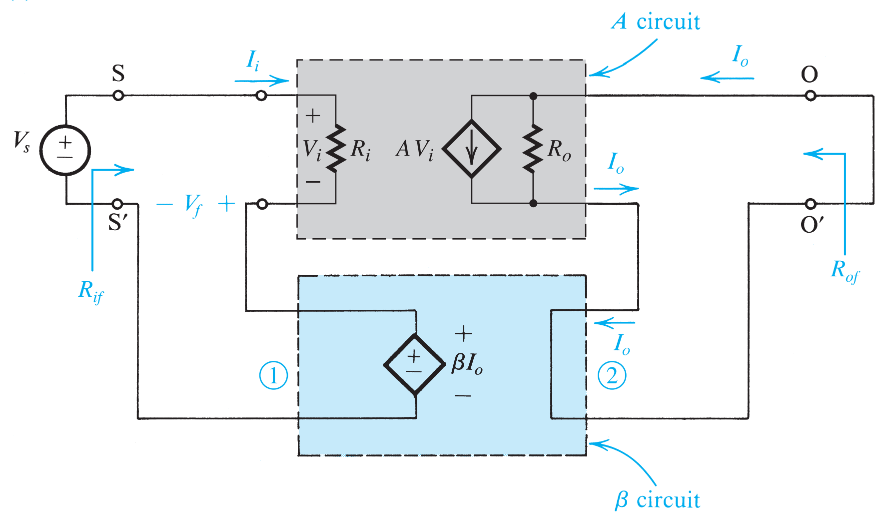
11.5.2 The Feedback Transconductance Amplifier (Series–Series)

Figure 11.18(a) shows the ideal structure of the feedback transconductance amplifier. The open-loop amplifier (A circuit) is unilateral, has an input resistance R_i, a short-circuit transconductance gain A in A/V, and an output resistance R_o. The short-circuit output current $I_o = AV_i$ is sensed by the feedback network. Note that since the resistance looking into port 2 of the feedback network is zero, the feedback network does not load the amplifier output. The feedback network provides at port 1 a voltage signal $V_f = \beta I_o$, where the feedback factor β has the dimensions of V/A. The feedback signal is connected in series with the input signal source V_s, and the feedback network does not load the amplifier input. Finally, note the definitions of the input resistance with feedback, R_{if}, and the output resistance with feedback, R_{of}. The latter is the resistance found by looking into the output loop between any two nodes such as O and O'.

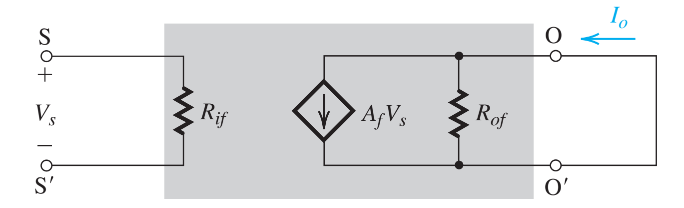
The equivalent circuit of the feedback transconductance amplifier is shown in Fig. 11.18(b). Note that the closed-loop gain A_f is the ratio of the short-circuit output current I_o and the input voltage V_s, thus it is the short-circuit transconductance of the feedback amplifier. Also, R_{of} is the resistance seen between any two nodes in the output loop, such as O and O', while V_s is set to zero. Finally, Fig. 11.18(c) provides the formulas for determining A_f, R_{if}, and R_{of}.

With the ideal case in hand, we now consider the general or practical case of a feedback transconductance amplifier, shown in Fig. 11.19(a). To be able to apply feedback analysis to this circuit, we have to find the A circuit and β. These are shown in Fig. 11.19(b). The A circuit is obtained by augmenting the basic amplifier with R_s and R_L and the two resistances R_{11} and R_{22}, which represent the loading effect of the feedback network on the basic amplifier at the input and output, respectively. Figure 11.19(b) shows how R_{11} and R_{22} are determined. Here, the series connection at both the input and the output means that the other port of the feedback network must be left open-circuited. This is also the case when β is determined.

(a) Ideal Structure



(b) Equivalent Circuit



(c) Formulas

$$A_f = \frac{I_o}{V_s} = \frac{A}{1 + A\beta}$$

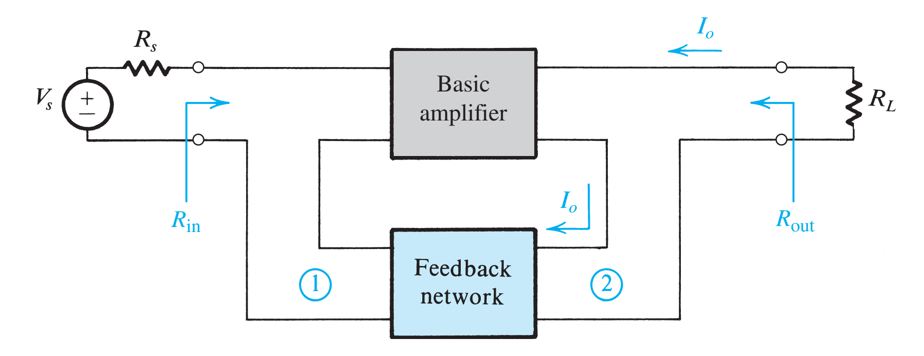
$$R_{if} = (1 + A\beta)R_i$$

$$R_{of} = (1 + A\beta)R_o$$

Figure 11.18 The feedback transconductance amplifier (series–series).

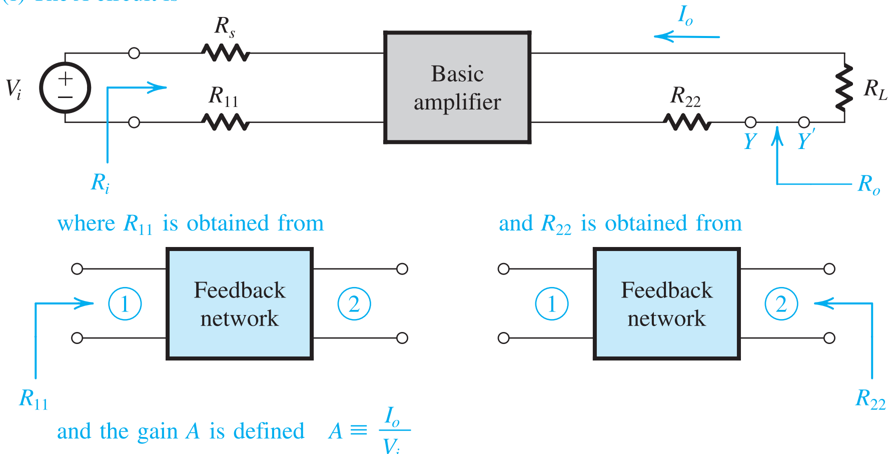
Finally Fig. 11.19(c) gives the formulas for determining the actual values of the input and output resistances, R_{in} and R_{out}, of the feedback amplifier from R_{if} and R_{of}. To see how these formulas come about, note from Fig. 11.19(a) that unlike R_{if}, R_{in} does not include R_s, and unlike R_{of}, R_{out} does not include R_L.

(a) General Structure

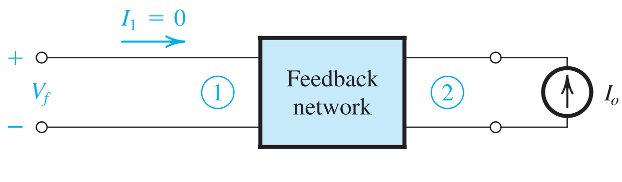


(b) Finding the A Circuit and β

(i) The A circuit is



(ii) β is obtained from



$$\beta \equiv \left.\frac{V_f}{I_o}\right|_{I_1 = 0}$$

(c) Gain, Input, and Output Resistance

- Use the formulas in Fig. 11.18 to find A_f, R_{if}, and R_{of}.
- R_{in} and R_{out} can then be found from

$$R_{in} = R_{if} - R_s$$
$$R_{out} = R_{of} - R_L$$

Figure 11.19 The feedback transconductance amplifier (series–series).

Example 11.7

Figure 11.20(a) shows a feedback transconductance amplifier composed of a differential amplifier A_1 with an input differential resistance R_{id}, an open-circuit voltage gain A_1, and an output resistance R_{o1}, connected in cascade with a common source MOSFET Q_2 having a transconductance g_{m2} and an output resistance r_{o2}. Use the feedback-analysis method to determine the closed-loop transconductance $A_f \equiv I_o/V_s$, the input resistance R_{in}, and the output resistance R_{out}. The latter is the resistance seen between the terminals of R_L, looking back into the output loop.

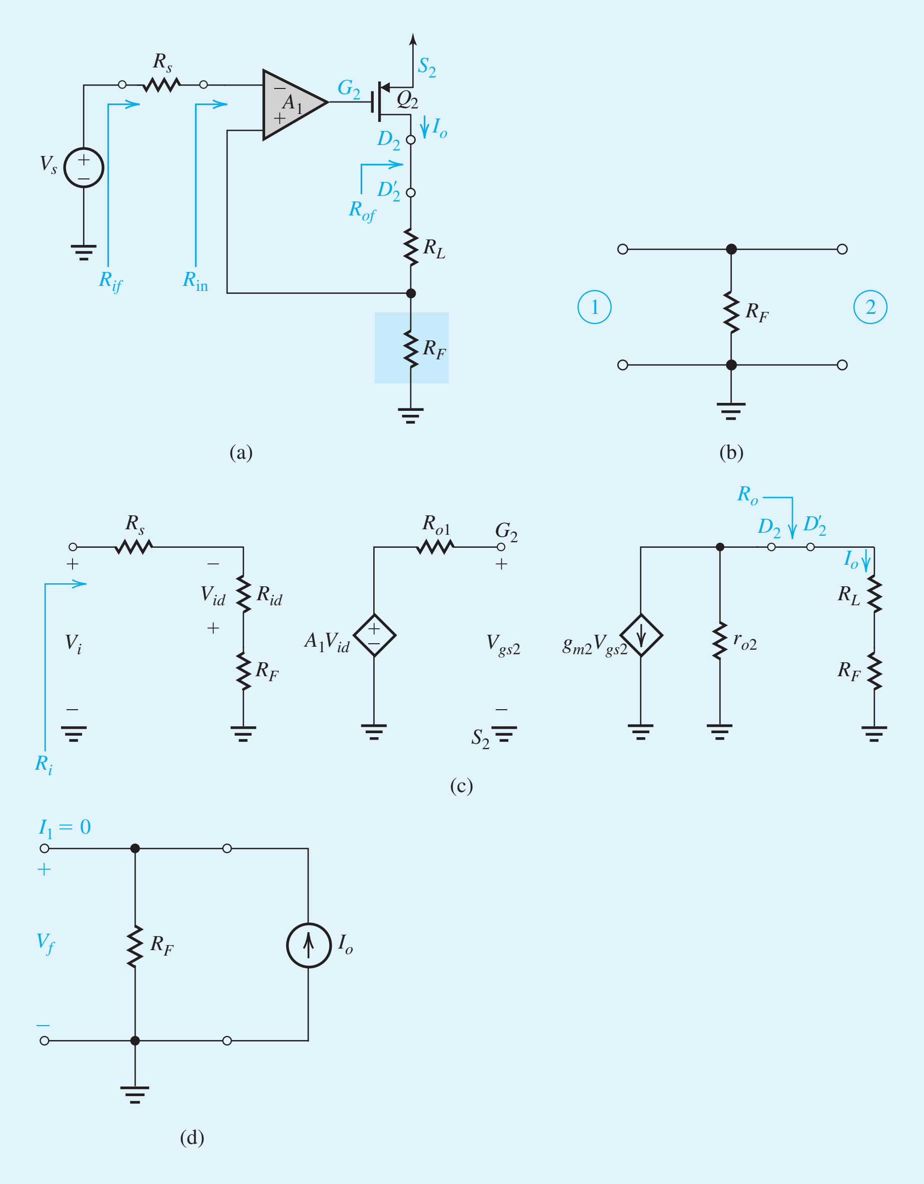


Figure 11.20 Circuits for Example 11.7.

Example 11.7 *continued*

Solution

First we identify the basic amplifier and the feedback circuit. The basic amplifier consists of the differential amplifier A_1 cascaded with the CS PMOS transistor Q_2. The output current I_o is sensed by the series resistance R_F. The latter is the feedback network (Fig. 11.20b). It develops a voltage V_f that is mixed in series with the input loop.

The second step is to ascertain that the feedback is negative. This can be done by assuming an increase in V_s and following the resulting change around the loop. An increase in V_s will cause the voltage at the inverting input terminal of A_1 to increase. This in turn causes a decrease in the output voltage of A_1, which is the voltage at the gate of Q_2. Thus, transistor Q_2 will have a larger V_{SG}, which will cause I_o to increase. This in turn causes an increase in the voltage across R_F, which is the same polarity as the change in V_s. Thus, the feedback is indeed negative.

Next, we determine an approximate value for $A_f \equiv I_o/V_s$ under the assumption that the loop gain $A\beta$ is much greater than unity. This value, found before any analysis is undertaken, will help us determine at the end whether our analysis is correct: If the loop gain is found to be much greater than unity, then the final A_f should be close to the value initially determined. From the circuit of Fig. 11.20(d),

$$\beta = R_F$$

and thus for large $A\beta$,

$$A_f \simeq \frac{1}{\beta} = \frac{1}{R_F}$$

Next, we determine the A circuit. Since the feedback network (Fig. 11.20b) is connected in series with both the input and output loops, we include a resistance R_F in each of these loops (which is equivalent to saying we include, at the input, the input resistance of the feedback circuit with port 2 open and, at the output, the input resistance of the feedback circuit with port 1 open). Doing this, including R_s and R_L in the A circuit, and replacing A_1 and Q_2 with their small-signal models, results in the A circuit shown in Fig. 11.20(c). Analysis of this circuit is straightforward:

$$V_{id} = -V_i \frac{R_{id}}{R_{id} + R_s + R_F} \tag{11.27}$$

$$V_{gs2} = A_1 V_{id} \tag{11.28}$$

$$I_o = -g_{m2} V_{gs2} \frac{r_{o2}}{r_{o2} + R_L + R_F} \tag{11.29}$$

Combining these three equations results in

$$A \equiv \frac{I_o}{V_i} = (A_1 g_{m2})\left(\frac{R_{id}}{R_{id} + R_s + R_F}\right)\left(\frac{r_{o2}}{r_{o2} + R_L + R_F}\right) \tag{11.30}$$

Usually $R_{id} \gg (R_s + R_F)$, $r_{o2} \gg (R_L + R_F)$, resulting in the approximate expression for A:

$$A \simeq A_1 g_{m2} \tag{11.31}$$

The input resistance R_i can be found by inspection as

$$R_i = R_s + R_{id} + R_F \tag{11.32}$$

The output resistance R_o is found by setting $V_i = 0$, and breaking the output loop at any location, say between D_2 and D_2'. Thus,

$$R_o = r_{o2} + R_L + R_F \tag{11.33}$$

The loop gain $A\beta$ is thus

$$A\beta = (A_1 g_{m2} R_F)\left(\frac{R_{id}}{R_{id} + R_s + R_F}\right)\left(\frac{r_{o2}}{r_{o2} + R_L + R_F}\right) \tag{11.34}$$

$$\simeq A_1 g_{m2} R_F \tag{11.35}$$

With numerical values, one can now obtain the value of $A\beta$ and determine whether it is indeed much greater than unity. We next determine the closed-loop gain

$$A_f = \frac{A}{1 + A\beta}$$

Substituting for A from Eq. (11.31) and for $A\beta$ from Eq. (11.35), we have

$$A_f \simeq \frac{A_1 g_{m2}}{1 + A_1 g_{m2} R_F} \tag{11.36}$$

For $A_1 g_{m2} R_F \gg 1$,

$$A_f \simeq \frac{1}{R_F}$$

which is the value we found at the outset.

The series mixing raises the input resistance with feedback,

$$R_{if} = R_i(1 + A\beta)$$

and R_{in} can be obtained by subtracting R_s from R_{if}.

To obtain R_{of}, we note that the series connection at the output raises the output resistance, thus,

$$R_{of} = R_o(1 + A\beta)$$

and R_{out}, which is the resistance seen by R_L, can be obtained by subtracting R_L from R_{of}.

EXERCISES

D11.12 For the circuit analyzed in Example 11.7, select a value for R_F that will result in $A_f \simeq 5$ mA/V. Now, for $A_1 = 200$ V/V, $g_{m2} = 2$ mA/V, $R_{id} = 100$ kΩ, $r_{o2} = 20$ kΩ, and assuming that $R_s \ll R_{id}$ and $R_L \ll r_{o2}$, find the value of A_f realized and the input and output resistances of the feedback transconductance amplifier. If for some reason g_{m2} drops in value by 50%, what is the corresponding percentage change in A_f?
Ans. 200 Ω; 4.94 mA/V; 8.1 MΩ; 1.64 MΩ; −1.25%

11.13 Determine the loop gain of the feedback amplifier of Fig. 11.20(a) by setting $V_s = 0$, breaking the feedback loop at G_2, applying a voltage V_{g2}, and determining the voltage V_{o1} that appears at the output of A_1; $A\beta \equiv -V_{o1}/V_{g2}$. Assume $R_F \ll (R_{id} + R_s)$ and $r_{o2} \gg R_L + R_F$.
Ans. $A\beta \simeq A_1 g_{m2} R_F$

11.14 Utilizing the full expression for $A\beta$ in Eq. (11.34) but assuming that $r_{o2} \gg R_L + R_F$ and $R_F \ll R_{id}$, show that

$$R_{\text{in}} \simeq R_{id}(1 + A_1 g_{m2} R_F)$$

11.15 Utilizing the full expression for $A\beta$ in Eq. (11.34) but assuming that $R_{id} \gg R_s + R_F$ and $R_F \ll r_{o2}$ show that

$$R_{\text{out}} \simeq r_{o2}(1 + A_1 g_{m2} R_F)$$

Example 11.8

Because negative feedback extends the amplifier bandwidth, it is commonly used in the design of broadband amplifiers. One such amplifier is the MC1553. Part of the circuit of the MC1553 is shown in Fig. 11.21(a). The circuit shown (called a **feedback triple**) is composed of three gain stages with series–series feedback provided by the network composed of R_{E1}, R_F, and R_{E2}.

Observe that the feedback network samples the emitter current I_o of Q_3, and thus I_o is the output quantity of the feedback amplifier. However, practically speaking, I_o is rather difficult to utilize. Thus it is usual to take as the output I_c, the collector current of Q_3. This current is of course almost equal to I_o; $I_c = \alpha I_o$. Thus, as a transconductance amplifier with I_c as the output current, the output resistance of interest is that labeled R_{out} in Fig. 11.21(a). In some applications, I_c is passed through a load resistance, such as R_{C3}, and the voltage V_o is taken as the output. Assume that the bias circuit, which is not shown, establishes $I_{C1} = 0.6$ mA, $I_{C2} = 1$ mA, and $I_{C3} = 4$ mA. Also assume that for all three transistors,[5] $h_{fe} = 100$ and $r_o = \infty$.

[5]To avoid possible confusion of the BJT current gain β and the feedback factor β, we sometimes use h_{fe} to denote the transistor β.

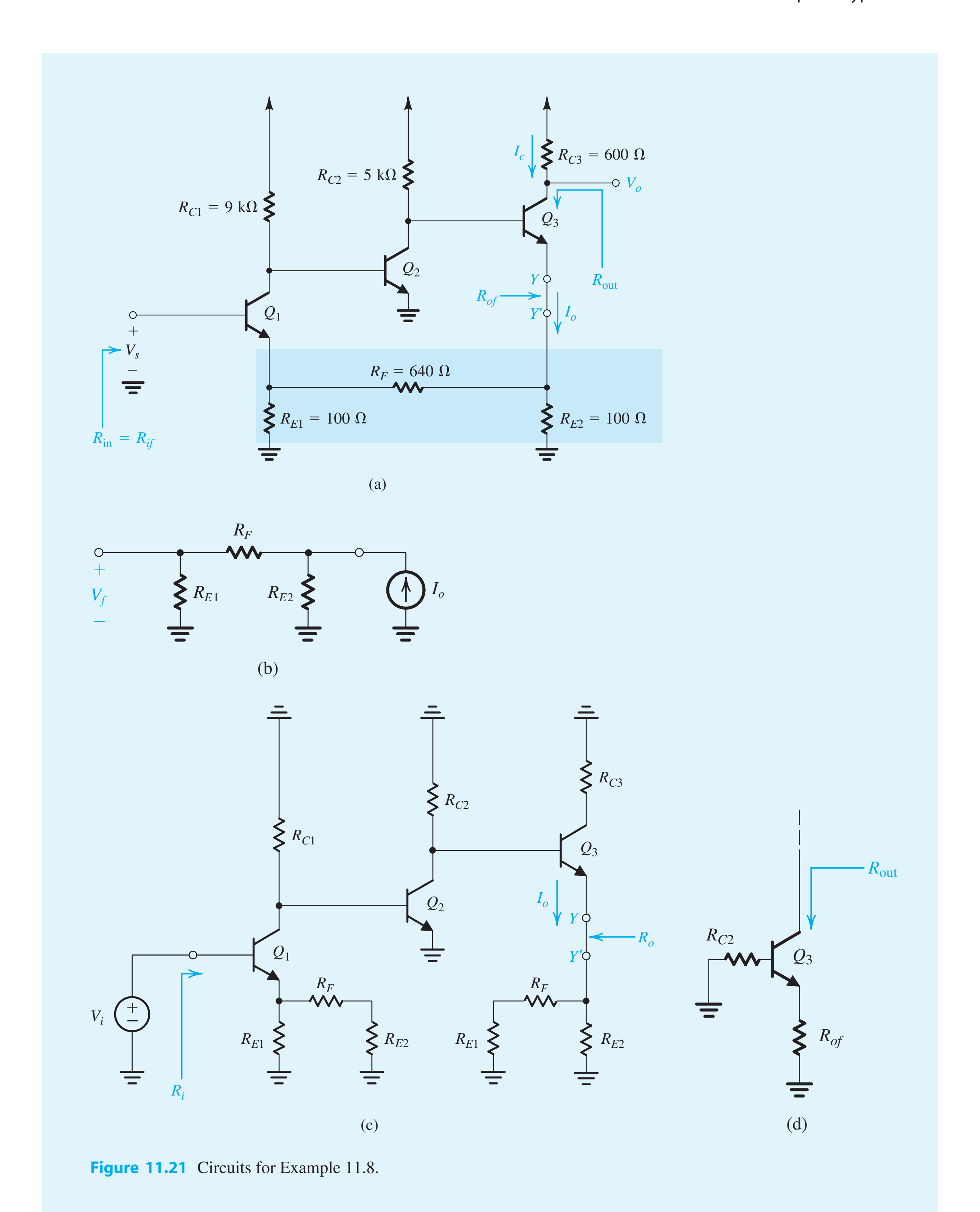


Figure 11.21 Circuits for Example 11.8.

Example 11.8 *continued*

(a) Anticipating that the loop gain will be large, find an approximate expression and value for the closed-loop gain $A_f \equiv I_o/V_s$ and hence for I_c/V_s. Also find V_o/V_s.
(b) Use feedback analysis to find A, β, A_f, V_o/V_s, R_{in}, and R_{out}. For the calculation of R_{out}, assume that r_o of Q_3 is 25 kΩ.

Solution

(a) When $A\beta \gg 1$,

$$A_f \equiv \frac{I_o}{V_s} \simeq \frac{1}{\beta}$$

where the feedback factor β can be found from the feedback network. The feedback network is highlighted in Fig. 11.21(a), and the determination of the value of β is illustrated in Fig. 11.21(b), from which we find

$$\beta \equiv \frac{V_f}{I_o} = \frac{R_{E2}}{R_{E2}+R_F+R_{E1}} \times R_{E1}$$
$$= \frac{100}{100+640+100} \times 100 = 11.9\ \Omega$$

Thus,

$$A_f \simeq \frac{1}{\beta}$$
$$= \frac{1}{R_{E2}}\left(1+\frac{R_{E2}+R_F}{R_{E1}}\right)$$
$$= \frac{1}{11.9} = 84\ \text{mA/V}$$
$$\frac{I_c}{V_s} \simeq \frac{I_o}{V_s} = 84\ \text{mA/V}$$
$$\frac{V_o}{V_s} = \frac{-I_c R_{C3}}{V_s} = -84 \times 0.6 = -50.4\ \text{V/V}$$

(b) Employing the loading rules given in Fig. 11.19, we obtain the A circuit shown in Fig. 11.21(c). To find $A \equiv I_o/V_i$ we first determine the gain of the first stage. This can be written by inspection as

$$\frac{V_{c1}}{V_i} = \frac{-\alpha_1 (R_{C1} \,\|\, r_{\pi 2})}{r_{e1} + [R_{E1} \,\|\, (R_F + R_{E2})]}$$

Since Q_1 is biased at 0.6 mA, $r_{e1} = 41.7\ \Omega$. Transistor Q_2 is biased as 1 mA; thus $r_{\pi 2} = h_{fe}/g_{m2} = 100/40 = 2.5\ \text{k}\Omega$. Substituting these values together with $\alpha_1 = 0.99$, $R_{C1} = 9\ \text{k}\Omega$, $R_{E1} = 100\ \Omega$, $R_F = 640\ \Omega$, and $R_{E2} = 100\ \Omega$, results in

$$\frac{V_{c1}}{V_i} = -14.92\ \text{V/V}$$

Next, we determine the gain of the second stage, which can be written by inspection (noting that $V_{b2} = V_{c1}$) as

$$\frac{V_{c2}}{V_{c1}} = -g_{m2}\{R_{C2} \parallel (h_{fe}+1)[r_{e3} + (R_{E2} \parallel (R_F + R_{E1}))]\}$$

Substituting $g_{m2} = 40\,\text{mA/V}$, $R_{C2} = 5\ \text{k}\Omega$, $h_{fe} = 100$, $r_{e3} = 25/4 = 6.25\ \Omega$, $R_{E2} = 100\,\Omega$, $R_F = 640\,\Omega$, and $R_{E1} = 100\,\Omega$ results in

$$\frac{V_{c2}}{V_{c1}} = -131.2\ \text{V/V}$$

Finally, for the third stage we can write by inspection

$$\frac{I_o}{V_{c2}} = \frac{I_{e3}}{V_{b3}} = \frac{1}{r_{e3} + (R_{E2} \parallel (R_F + R_{E1}))}$$

$$= \frac{1}{6.25 + (100 \parallel 740)} = 10.6\ \text{mA/V}$$

Combining the gains of the three stages results in

$$A \equiv \frac{I_o}{V_i} = -14.92 \times -131.2 \times 10.6 \times 10^{-3}$$

$$= 20.7\ \text{A/V}$$

The closed-loop gain A_f can now be found from

$$A_f \equiv \frac{I_o}{V_s} = \frac{A}{1 + A\beta}$$

$$= \frac{20.7}{1 + 20.7 \times 11.9} = 83.7\ \text{mA/V}$$

which we note is very close to the ideal value found in (a) above. This is not surprising, since the loop gain $A\beta = 20.7 \times 11.9 = 246.3$ is large ($\gg 1$).

The voltage gain is found from

$$\frac{V_o}{V_s} = \frac{-I_c R_{C3}}{V_s} \simeq \frac{-I_o R_{C3}}{V_s} = -A_f R_{C3}$$

$$= -83.7 \times 10^{-3} \times 600 = -50.2\ \text{V/V}$$

which is also very close to the approximate value found in (a) above.

Example 11.8 *continued*

The input resistance of the feedback amplifier is given by

$$R_{in} = R_{if} = R_i(1 + A\beta)$$

where R_i is the input resistance of the A circuit. The value of R_i can be found from the circuit in Fig. 11.21(c) as follows:

$$\begin{aligned} R_i &= (h_{fe} + 1)\left[r_{e1} + (R_{E1} \| (R_F + R_{E2}))\right] \\ &= 13.11 \text{ k}\Omega \end{aligned}$$

Thus,

$$R_{if} = 13.11(1 + 20.7 \times 11.9) = 3.24 \text{ M}\Omega$$

To determine the output resistance R_{out}, which is the resistance looking into the collector of Q_3, we face a dilemma. The feedback does not sample I_c and thus we cannot employ the feedback formulas directly.[6] Nevertheless, we present a somewhat indirect solution to this problem below. Here we note parenthetically that had Q_1 been a MOSFET, this problem would not have existed, since $I_d = I_s$.

Since the feedback senses the emitter current I_o, the output resistance given by the feedback analysis will be the resistance seen in the emitter circuit, say between Y and Y',

$$R_{of} = R_o(1 + A\beta)$$

where R_o can be determined from the A circuit in Fig. 11.21(c) by breaking the circuit between Y and Y'. The resistance looking between these two nodes can be found to be

$$R_o = \left[R_{E2} \| (R_F + R_{E1})\right] + r_{e3} + \frac{R_{C2}}{h_{fe} + 1}$$

which, for the values given, yields $R_o = 143.9\ \Omega$. The output resistance R_{of} of the feedback amplifier can now be found as

$$R_{of} = R_o(1 + A\beta) = 143.9(1 + 20.7 \times 11.9) = 35.6 \text{ k}\Omega$$

[6]This important point was first brought to the authors' attention by Gordon Roberts (see Roberts and Sedra, 1992).

We can now use the value of R_{of} to obtain an approximate value for R_{out}. To do this, we assume that the effect of the feedback is to place a resistance R_{of} (35.6 kΩ) in the emitter of Q_3, and find the output resistance from the equivalent circuit shown in Fig. 11.21(d). This is the output resistance of a BJT with a resistance R_{of} in its emitter and a resistance R_{C2} in its base. The formula we have for this (Eq. 8.66) does not unfortunately account for a resistance in the base. The formula, however, can be modified (see Problem 11.54) to obtain

$$R_{out} = r_{o3} + \left[R_{of} \parallel \left(r_{\pi 3} + R_{C2}\right)\right]\left[1 + g_{m3} r_{o3} \frac{r_{\pi 3}}{r_{\pi 3} + R_{C2}}\right]$$

$$= 25 + [35.6 \parallel (0.625 + 5)]\left[1 + 160 \times 25 \times \frac{0.625}{0.625 + 5}\right]$$

$$= 2.19\,\text{M}\Omega$$

Thus R_{out} is increased (from r_{o3}) but not by $(1 + A\beta)$.

EXERCISES

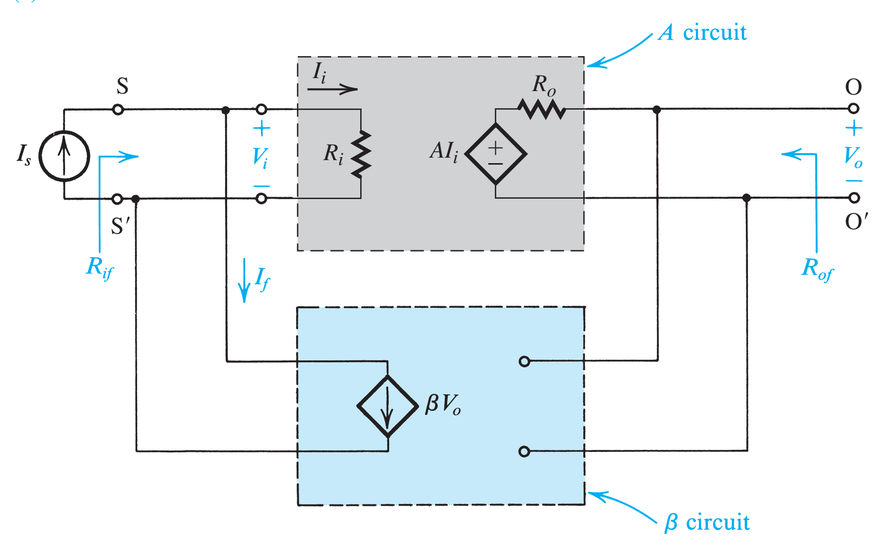
D11.16 For the feedback triple in Fig. 11.21(a), analyzed in Example 11.8, modify the value of R_F to obtain a closed-loop transconductance I_o/V_s of approximately 100 mA/V. Assume that the loop gain remains large. What is the new value of R_F? For this value, what is the approximate value of the voltage gain if the output voltage is taken at the collector of Q_3?
Ans. 800 Ω; –60 V/V

11.17 Determine the loop gain of the feedback amplifier of Fig. 11.21(a). Set $V_s = 0$, break the loop between the collector of Q_1 and the base of Q_2, apply a voltage V_t to the base of Q_2, and connect a resistance equal to $r_{\pi 2}$ between the collector of Q_1 and ground. Find $A\beta$ as $(-V_r/V_t)$ where $V_r = V_{c1}$.
Ans. $A\beta = 248.9$ (slightly different from the value found in Example 11.8 because of the approximations inherent in the systematic feedback-analysis method).

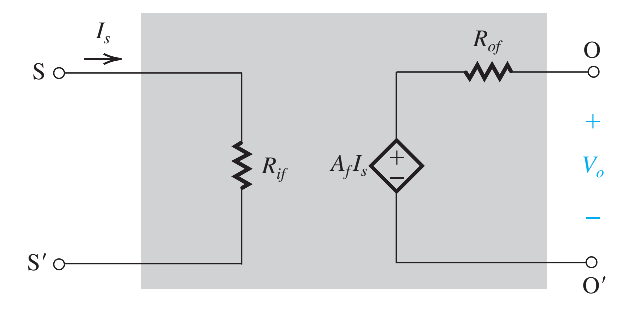
11.5.3 The Feedback Transresistance Amplifier (Shunt–Shunt)

Figure 11.22(a) shows the ideal feedback transresistance amplifier, which, as expected, utilizes the shunt–shunt topology. The amplifier equivalent circuit is shown in Fig. 11.22(b), with the formulas for determining A_f, R_{if}, and R_{of} given in Fig. 11.22(c).

(a) Ideal Structure



(b) Equivalent Circuit



(c) Formulas

$$A_f \equiv \frac{V_o}{I_s} = \frac{A}{1 + A\beta}$$

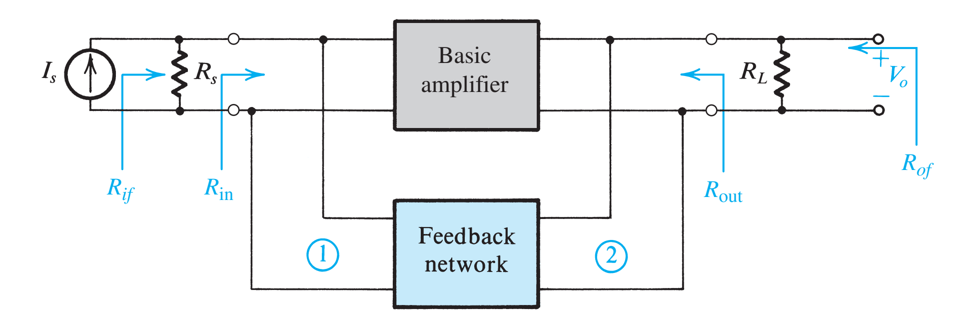
$$R_{if} = R_i/(1 + A\beta)$$

$$R_{of} = R_o/(1 + A\beta)$$

Figure 11.22 The feedback transresistance amplifier (shunt–shunt).

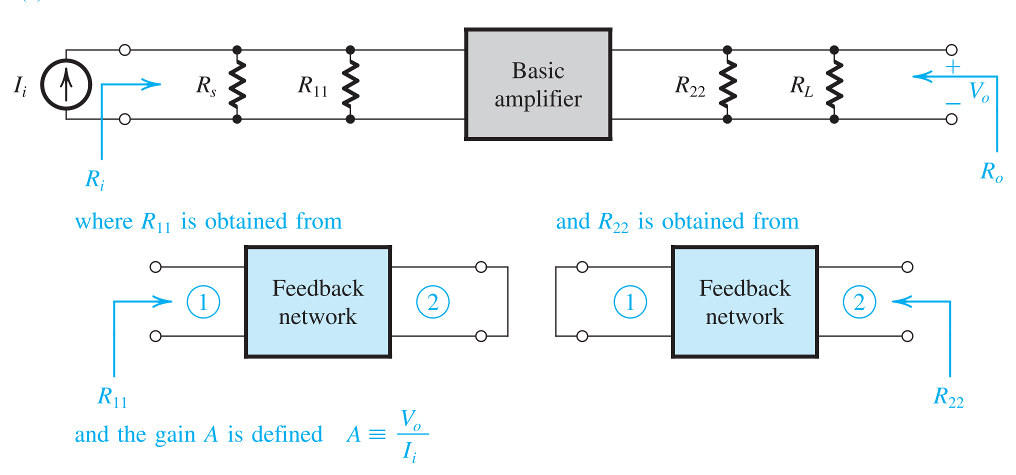
Application of the feedback-analysis method to a general transresistance amplifier is shown in Fig. 11.23, which presents all the steps and formulas needed. The method is now illustrated by a detailed example.

(a) General Structure

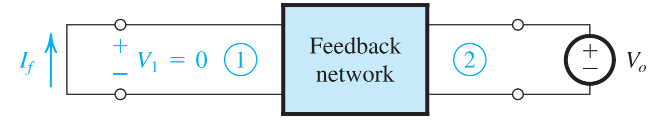


(b) Finding the A Circuit and β

(i) The A circuit is



(ii) β is obtained from



$$\beta \equiv \left.\frac{I_f}{V_o}\right|_{V_1 = 0}$$

(c) Gain, Input, and Output Resistance

- Use the formulas in Fig. 11.21 to find A_f, R_{if}, and R_{of}.
- R_{in} and R_{out} can then be found from

$$R_{in} = 1\bigg/\left(\frac{1}{R_{if}} - \frac{1}{R_s}\right)$$

$$R_{out} = 1\bigg/\left(\frac{1}{R_{of}} - \frac{1}{R_L}\right)$$

Figure 11.23 The feedback transresistance amplifier (shunt–shunt).

Example 11.9

Figure 11.24(a) shows a feedback transresistance amplifier. It is formed by connecting a resistance R_F in the negative-feedback path of a voltage amplifier with gain μ, an input resistance R_{id}, and an output resistance r_o. The amplifier μ can be implemented with an op amp, a simple differential amplifier, a single-ended inverting amplifier, or, in the limit, a single-transistor CE or CS amplifier. The latter case will be considered in Exercise 11.18. Of course, the higher the gain μ, the more ideal the characteristics of the feedback transresistance amplifier will be, simply because of the concomitant increase in loop gain.

(a) If the loop gain is large, find an approximate expression for the closed-loop transresistance V_o/I_s of the feedback amplifier.
(b) Find the A circuit and expressions for A, R_i, and R_o.
(c) Find expressions for the loop gain, A_f, R_{if}, R_{in}, R_{of}, and R_{out}.
(d) Find the values of R_i, R_o, A, β, $A\beta$, A_f, R_{if}, R_{in}, R_{of}, and R_{out} for the case $\mu = 10^4$ V/V, $R_{id} = \infty$, $r_o = 100\ \Omega$, $R_F = 10\ \text{k}\Omega$, and $R_s = R_L = 1\ \text{k}\Omega$.
(e) If instead of a current source I_s having a source resistance $R_s = 1\ \text{k}\Omega$, the amplifier is fed from a voltage source V_s having a source resistance $R_s = 1\ \text{k}\Omega$, find an expression for and the value of the voltage gain V_o/V_s.

Solution

(a) If the loop gain $A\beta$ is large,

$$A_f \equiv \frac{V_o}{I_s} \simeq \frac{1}{\beta}$$

where β can be found from the β circuit in Fig. 11.24(b) as shown in Fig. 11.24(c),

$$\beta \equiv \frac{I_f}{V_o} = -\frac{1}{R_F} \tag{11.37}$$

Thus,

$$\frac{V_o}{I_s} \simeq -R_F$$

Note that in this case the voltage at the input node (the inverting input terminal of μ) will be very close to ground and thus very little, if any, current flows into the input terminal of the amplifier. Nearly all of I_s will flow through R_F, resulting in $V_o \simeq 0 - I_s R_F = -I_s R_F$. This should be reminiscent of the inverting op-amp configuration studied in Section 2.2.

(b) From the feedback network in Fig. 11.24(b), we see that the loading effect at the amplifier input and output will simply be R_F. This is indicated in the A circuit shown in Fig. 11.24(c), where we have replaced the amplifier μ with its equivalent-circuit model. The open-loop transresistance A can be obtained as follows:

$$V_{id} = I_i R_i \tag{11.38}$$

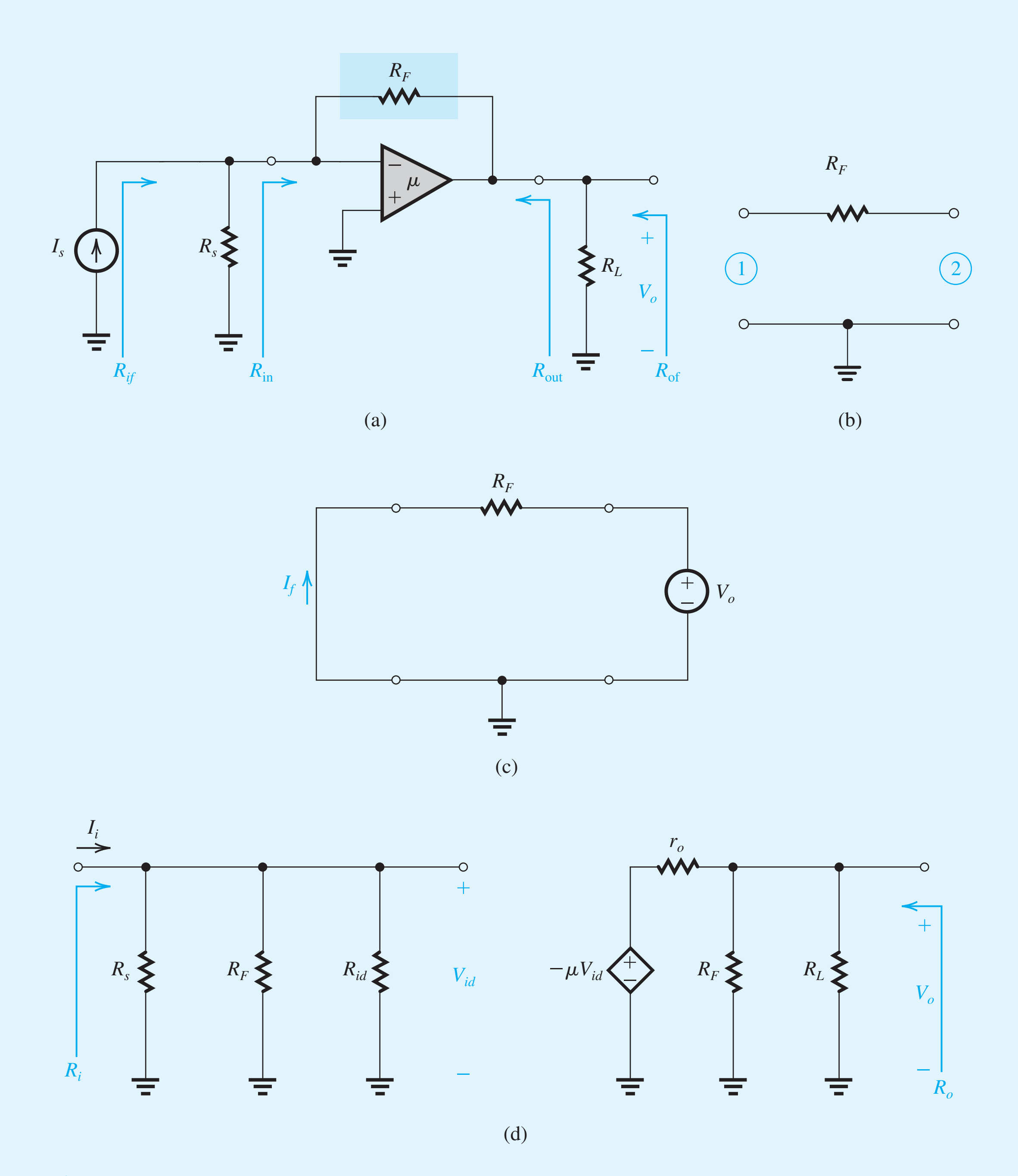


Figure 11.24 (**a**) A feedback transresistance amplifier; (**b**) the β circuit; (**c**) determining β; (**d**) the A circuit.

Example 11.9 *continued*

where

$$R_i = R_{id} \parallel R_F \parallel R_s \tag{11.39}$$

$$V_o = -\mu V_{id} \frac{(R_F \parallel R_L)}{r_o + (R_F \parallel R_L)} \tag{11.40}$$

Combining Eqs. (11.38) and (11.40) gives

$$A \equiv \frac{V_o}{I_i} = -\mu R_i \frac{(R_F \parallel R_L)}{r_o + (R_F \parallel R_L)} \tag{11.41}$$

The open-loop output resistance can be obtained by inspection of the A circuit with I_i set to 0. We see that $V_{id} = 0$, and

$$R_o = r_o \parallel R_F \parallel R_L \tag{11.42}$$

(c) The loop gain $A\beta$ can be obtained by combining Eqs. (11.37) and (11.41),

$$A\beta = \mu\left(\frac{R_i}{R_F}\right)\frac{(R_F \parallel R_L)}{r_o + (R_F \parallel R_L)} \tag{11.43}$$

Observe that although both A and β are negative, $A\beta$ is positive, a comforting fact confirming that the feedback is negative. Also note that $A\beta$ is dimensionless, as it must always be.

The closed-loop gain A_f can now be found as

$$A_f \equiv \frac{V_o}{I_s} = \frac{A}{1 + A\beta}$$

Thus

$$A_f = \frac{-\mu R_i \dfrac{(R_F \parallel R_L)}{r_o + (R_F \parallel R_L)}}{1 + \mu \dfrac{R_i}{R_F} \dfrac{(R_F \parallel R_L)}{r_o + (R_F \parallel R_L)}} \tag{11.44}$$

Note that the condition of $A\beta \gg 1$ that results in $A_f \simeq -R_F$ corresponds to

$$\mu\left(\frac{R_i}{R_F}\right)\frac{(R_F \parallel R_L)}{r_o + (R_F \parallel R_L)} \gg 1 \tag{11.45}$$

The input resistance with feedback, R_{if}, is obtained by dividing R_i by $(1+A\beta)$ with the result

$$R_{if} = \frac{R_i}{1+A\beta}$$

or

$$\frac{1}{R_{if}} = \frac{1}{R_i} + \frac{A\beta}{R_i} = \frac{1}{R_i} + \frac{\mu}{R_F}\frac{(R_F \parallel R_L)}{r_o + (R_F \parallel R_L)}$$

Substituting for R_i from Eq. (11.39) and replacing $\mu(R_F \parallel R_L)/[r_o + (R_F \parallel R_L)]$ by μ', where μ' is lower than but usually close to the value of μ, results in

$$R_{if} = R_{id} \parallel R_F \parallel R_s \parallel (R_F/\mu')$$

The two terms containing R_F can be combined,

$$R_{if} = R_s \parallel R_{id} \parallel [R_F/(\mu'+1)] \tag{11.46}$$

Since $R_{if} = R_s \parallel R_{\text{in}}$, we see that

$$R_{\text{in}} = R_{id} \parallel [R_F/(\mu'+1)]$$

Usually R_{id} is large and thus

$$R_{\text{in}} \simeq \frac{R_F}{\mu'+1} \simeq \frac{R_F}{\mu'} \tag{11.47}$$

from which we observe that for large amplifier gain μ, the input resistance will be low.

The output resistance with feedback R_{of} can be found by dividing R_o by $(1+A\beta)$:

$$R_{of} = \frac{R_o}{1+A\beta}$$

Thus,

$$\frac{1}{R_{of}} = \frac{1}{R_o} + \frac{A\beta}{R_o}$$

$$= \frac{1}{R_o} + \mu\frac{R_i}{R_F}\frac{(R_F \parallel R_L)}{r_o + (R_F \parallel R_L)}\frac{1}{R_o}$$

Example 11.9 *continued*

Substituting for R_o from Eq. (11.42),

$$\frac{1}{R_{of}} = \frac{1}{R_L} + \frac{1}{R_F} + \frac{1}{r_o} + \mu\frac{R_i}{R_F}\frac{1}{r_o}$$

$$= \frac{1}{R_L} + \frac{1}{R_F} + \frac{1}{r_o}\left(1 + \mu\frac{R_i}{R_F}\right)$$

Thus,

$$R_{of} = R_L \parallel R_F \parallel \frac{r_o}{1 + \mu\dfrac{R_i}{R_F}}$$

Since, moreover,

$$R_{of} = R_L \parallel R_{\text{out}}$$

we obtain for R_{out}

$$R_{\text{out}} = R_F \parallel \frac{r_o}{1 + \mu\dfrac{R_i}{R_F}}$$

Usually $R_F \gg r_o / \left[1 + \mu\left(R_i/R_F\right)\right]$; thus,

$$R_{\text{out}} \simeq \frac{r_o}{1 + \mu\dfrac{R_i}{R_F}} \simeq \left(\frac{R_F}{R_i}\right)\left(\frac{r_o}{\mu}\right)$$

from which we see that for large μ, the output resistance will be considerably reduced.
(d) For the numerical values given:

$$R_i = R_{id} \parallel R_F \parallel R_s$$

$$= \infty \parallel 10 \parallel 1 = 0.91 \text{ k}\Omega$$

$$R_o = r_o \parallel R_F \parallel R_s$$

$$= 0.1 \parallel 10 \parallel 1 = 90\ \Omega$$

$$A = -\mu R_i \frac{\left(R_F \parallel R_L\right)}{r_o + \left(R_F \parallel R_L\right)}$$

$$= -10^4 \times 0.91 \times \frac{(10 \parallel 1)}{0.1 + (10 \parallel 1)} = -8198 \text{ k}\Omega$$

$$\beta = -\frac{1}{R_F} = -\frac{1}{10} = -0.1 \text{ mA/V}$$

$$A\beta = 819.8$$

$$1 + A\beta = 820.8$$

$$A_f = \frac{A}{1 + A\beta} = -\frac{8198}{820.8} = -9.99 \text{ k}\Omega$$

which is very close to the ideal value of $-R_F = -10$ kΩ.

$$R_{if} = \frac{R_i}{1 + A\beta} = \frac{910}{820.8} = 1.11 \ \Omega$$

$$R_{\text{in}} = \frac{1}{\dfrac{1}{R_{if}} - \dfrac{1}{R_i}} - \frac{1}{\dfrac{1}{1.11} - \dfrac{1}{1000}} \simeq 1.11 \ \Omega$$

which is very low, a highly desirable property. We also have

$$R_{of} = \frac{R_o}{1 + A\beta} = \frac{90}{820.8} = 0.11 \ \Omega$$

$$R_{\text{out}} = \frac{1}{\dfrac{1}{R_{of}} - \dfrac{1}{R_L}} = \frac{1}{\dfrac{1}{0.11} - \dfrac{1}{1000}} \simeq 0.11 \ \Omega$$

which as well is very low, another highly desirable property.

(e) If the amplifier is fed with a voltage source V_s having a resistance $R_s = 1$ kΩ, the output voltage can be found from

$$V_o = A_f I_s = A_f \frac{V_s}{R_s}$$

Thus,

$$\frac{V_o}{V_s} = \frac{A_f}{R_s} = \frac{9.99 \text{ k}\Omega}{1 \text{ k}\Omega} = -9.99 \text{ V/V}$$

EXERCISES

11.18 It is required to determine the loop gain of the amplifier of Fig. 11.24(a) directly. Set $I_s = 0$, replace the amplifier μ with its equivalent circuit, and break the loop at the amplifier input, ensuring that a resistance equal to R_{id} is connected across R_s. Show that

$$A\beta = \frac{\mu R_L (R_{id} \| R_s)}{r_o[R_L + R_F + (R_{id} \| R_s)] + R_L[R_F + (R_{id} \| R_s)]}$$

Evaluate $A\beta$ using the numerical values given in Example 11.9.

Ans. $A\beta = 819.7$

11.19 For the transresistance amplifier in Fig. E11.19, replace the MOSFET with its small-signal equivalent-circuit model and use feedback analysis to show the following:

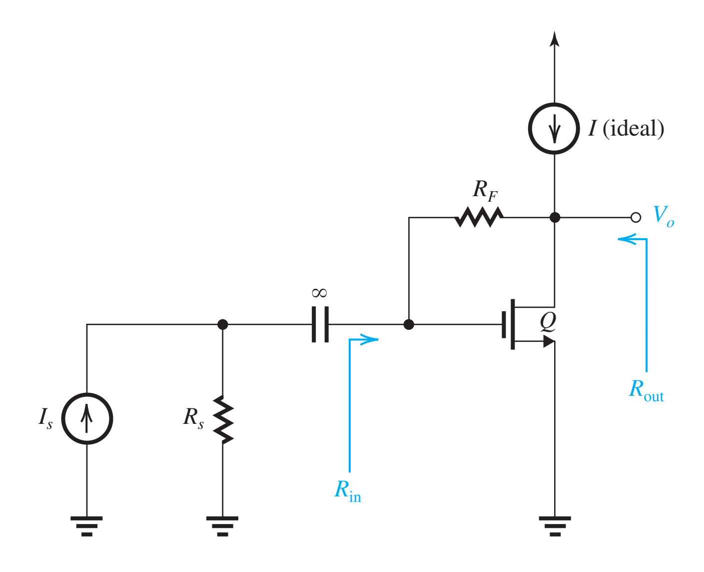


Figure E11.19

(a) For large loop gain (which cannot be achieved here), $A_f \equiv V_o/I_s \simeq -R_F$.

(b) $A_f = \dfrac{-(R_s \| R_F) g_m (r_o \| R_F)}{1 + (R_s \| R_F) g_m (r_o \| R_F)/R_F}$

(c) $R_{in} = \dfrac{R_F}{[1 + g_m (r_o \| R_F)]}$

(d) $R_{out} = r_o \| \dfrac{R_F}{1 + g_m (R_s \| R_F)}$

(e) For $g_m = 5$ mA/V, $r_o = 20$ kΩ, $R_F = 10$ kΩ, and $R_s = 1$ kΩ, find A, β, $A\beta$, A_f R_i, R_o, R_{if}, R_{in}, R_{of}, and R_{out}.

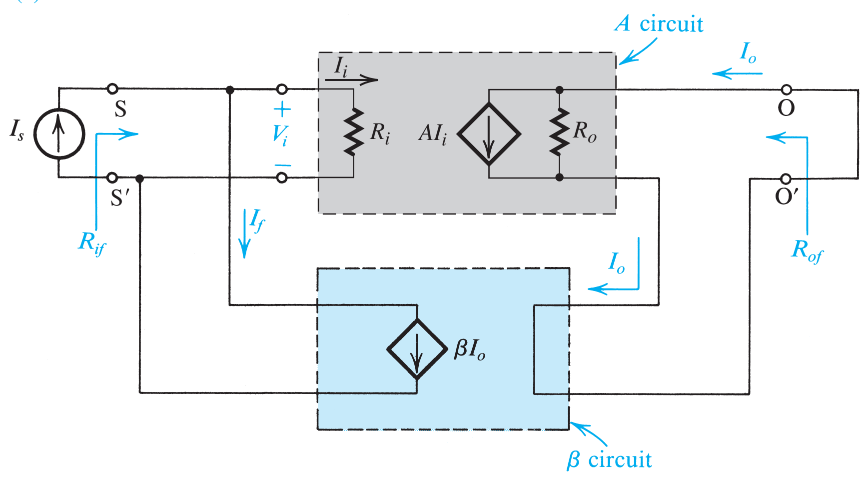
Ans. (e) −30.3 kΩ; −0.1 mA/V; 3.03; −7.52 kΩ (compare to the ideal value of −10 kΩ); 909 Ω; 6.67 kΩ; 226 Ω; 291 Ω; 1.66 kΩ; 1.66 kΩ

11.5.4 The Feedback Current Amplifier (Shunt–Series)

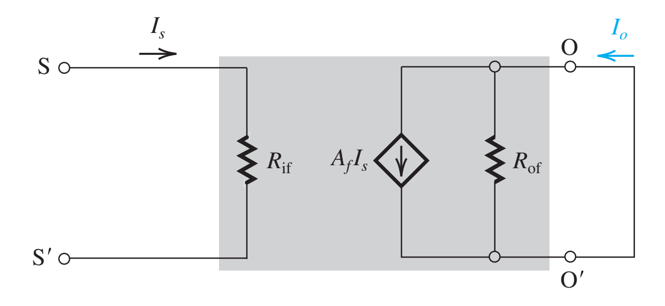
Figure 11.25(a) shows the ideal feedback current amplifier, which, as expected, utilizes the shunt–series topology. The amplifier equivalent circuit is shown in Fig. 11.25(b), with the formulas for determining A_f, R_{if}, and R_{of} given in Fig. 11.25(c).

Figure 11.26 shows how the feedback-analysis method can be applied to a general feedback current amplifier. All the necessary analysis steps and the formulas are given. The method is now illustrated by a detailed example.

(a) Ideal Structure



(b) Equivalent Circuit



(c) Formulas

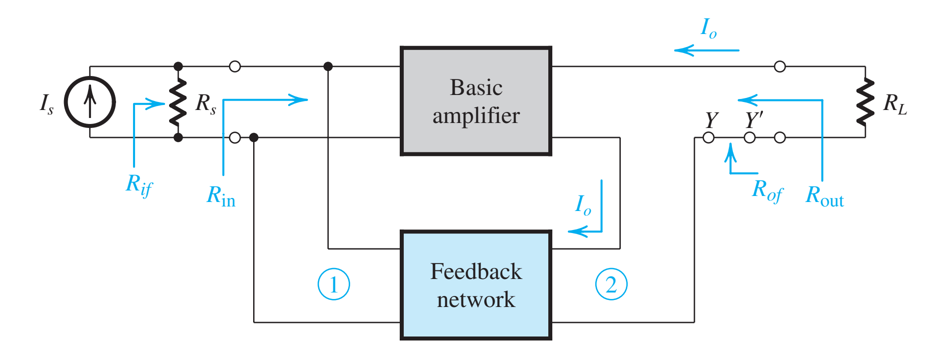
$$A_f \equiv \frac{I_o}{I_s} = \frac{A}{1 + A\beta}$$

$$R_{if} = R_i/(1 + A\beta)$$

$$R_{of} = (1 + A\beta)R_o$$

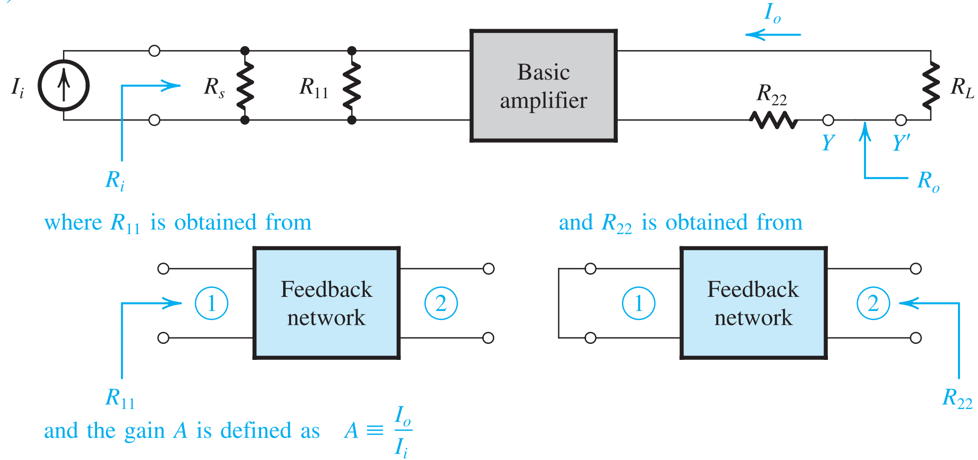
Figure 11.25 The feedback current amplifier (shunt–series).

(a) General Structure

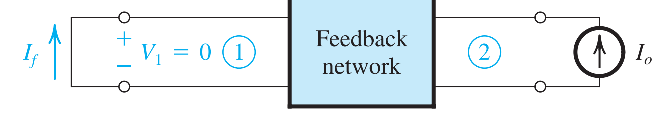


(b) Finding the A Circuit and β

(i) The A circuit is



(ii) β is obtained from



$$\beta \equiv \left.\frac{I_f}{I_o}\right|_{V_1 = 0}$$

(c) Gain, Input, and Output Resistance

- Use the formulas in Fig. 11.25 to find A_f, R_{if}, and R_{of}.
- R_{in} and R_{out} can then be found from

$$R_{in} = 1 \bigg/ \left(\frac{1}{R_{if}} - \frac{1}{R_s}\right)$$

$$R_{out} = R_{of} - R_L$$

Figure 11.26 The feedback current amplifier (shunt–series).

Example 11.10

Figure 11.27 shows a feedback current amplifier formed by cascading an inverting voltage amplifier μ with a MOSFET Q. The output current I_o is the drain current of Q. The feedback network, consisting of resistors R_1 and R_2, senses an exactly equal current, namely, the source current of Q, and provides a feedback current signal that is mixed with I_s at the input node. Note that the bias arrangement is *not* shown.

The amplifier μ can be implemented in a variety of ways, including by means of an op amp, a differential amplifier, or a single-ended inverting amplifier. The simplest approach is to implement μ with a CS

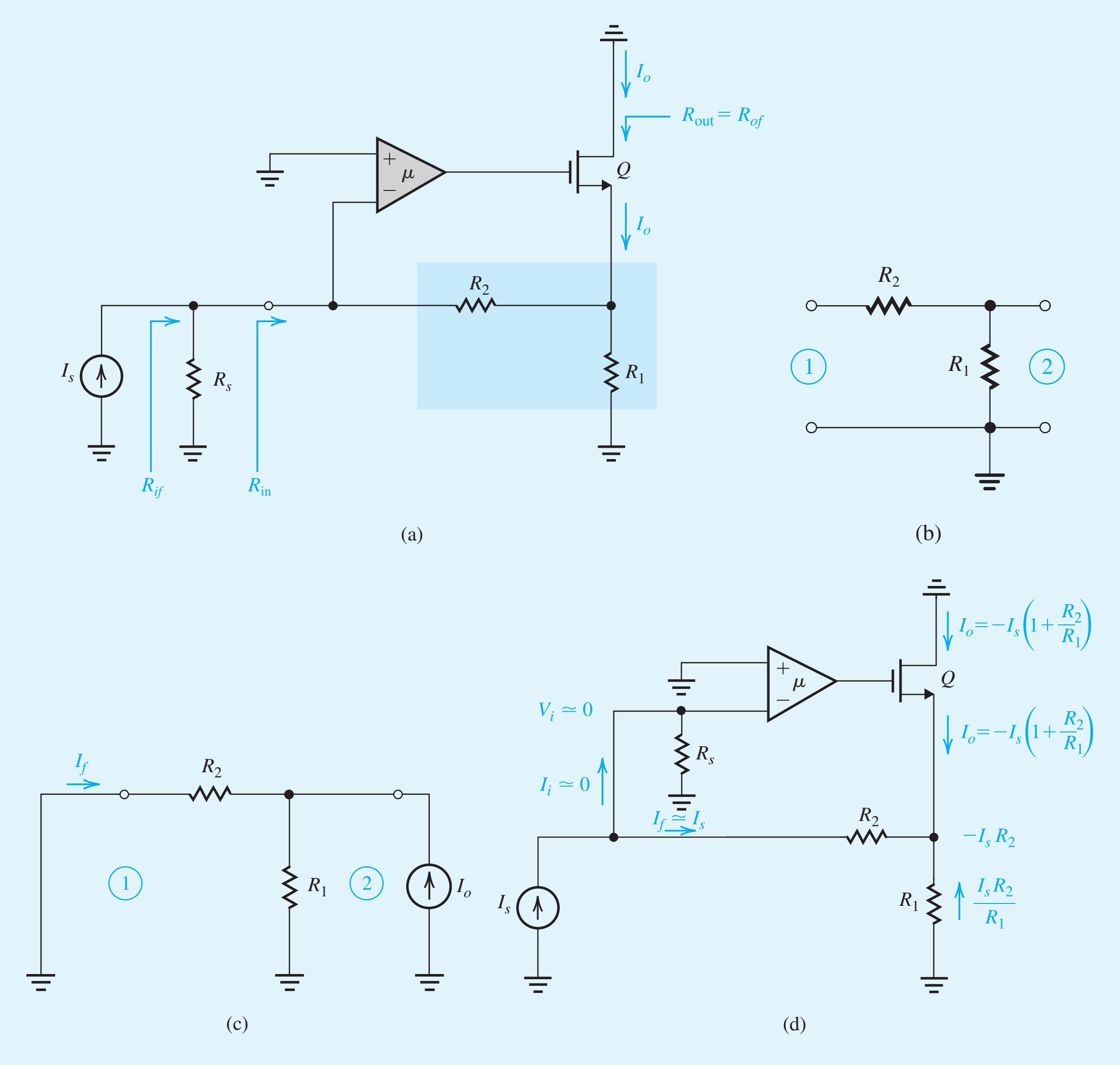


Figure 11.27 Circuit for Example 11.10.

Example 11.10 *continued*

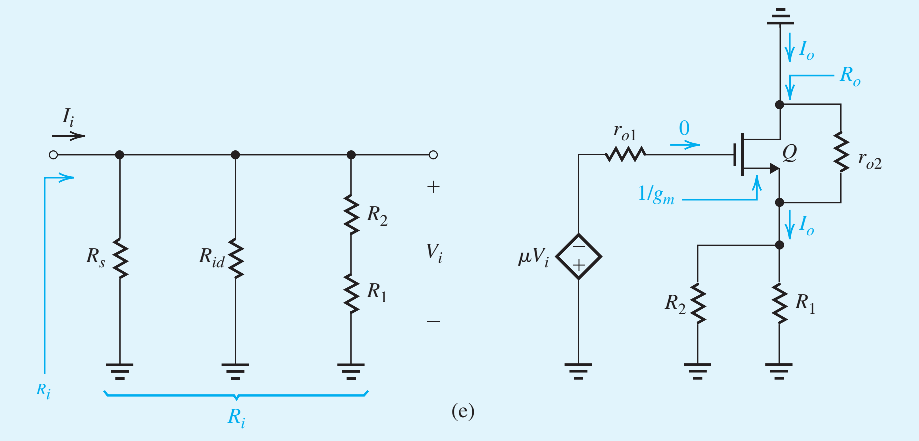


Figure 11.27 *continued*

MOSFET amplifier. However, in such a case the loop gain will be very limited. Assume that the amplifier μ has an input resistance R_{id}, an open-circuit voltage gain μ, and an output resistance r_{o1}.

(a) If the loop gain is large, find an approximate expression for the closed-loop gain $A_f \equiv I_o/I_s$.
(b) Find the A circuit and derive expressions for A, R_i, and R_o.
(c) Give expressions for $A\beta$, A_f, R_{if}, R_{in}, R_{of}, and R_{out}.
(d) Find numerical values for A, β, $A\beta$, A_f, R_i, R_{if}, R_{in}, R_o, R_{of}, and R_{out} for the following case: $\mu = 1000$ V/V, $R_s = \infty$, $R_{id} = \infty$, $r_{o1} = 1$ kΩ, $R_1 = 10$ kΩ, $R_2 = 90$ kΩ; for Q: $g_m = 5$ mA/V and $r_o = 20$ kΩ.

Solution

(a) When the loop gain $A\beta \gg 1$, $A_f \simeq 1/\beta$. The β circuit, shown in Fig. 11.27(b), is used to determine β as shown in Fig. 11.27(c),

$$\beta \equiv \frac{I_f}{I_o} = -\frac{R_1}{R_1 + R_2} \tag{11.48}$$

Thus,

$$A_f \simeq \frac{1}{\beta} = -\left(1 + \frac{R_2}{R_1}\right) \tag{11.49}$$

To see what happens in this case more clearly, refer to Fig. 11.27(d). Here we have assumed the loop gain to be large, so that $I_i \simeq 0$ and thus $I_f \simeq I_s$. Also note that because $I_i \simeq 0$, V_i will be close to zero. Thus, we can easily determine the voltage at the source of Q as $-I_f R_2 \simeq -I_s R_2$. The current through R_1 will then be $I_s R_2/R_1$. The source current of Q will be $-(I_s + I_s R_2/R_1)$, which means that the output current I_o will be

$$I_o = -I_s\left(1 + \frac{R_2}{R_1}\right)$$

which confirms the expression for A_f obtained above (Eq. 11.49).

(b) To obtain the A circuit we load the input side of the basic amplifier with R_s and R_{11}. The latter in this case is simply $R_1 + R_2$ (because port 2 of the feedback network is opened). We also load the output of the basic amplifier with R_{22}, which in this case is $R_1 \parallel R_2$ (because port 1 of the feedback network is shorted). The resulting A circuit is shown in Fig. 11.27(e), where we have replaced the amplifier μ with its equivalent circuit. Analysis of the A circuit is straightforward and proceeds as follows:

$$R_i = R_s \parallel R_{id} \parallel (R_1 + R_2) \tag{11.50}$$

$$V_i = I_i R_i \tag{11.51}$$

$$I_o = -\mu V_i \frac{1}{1/g_m + (R_1 \parallel R_2 \parallel r_{o2})} \frac{r_{o2}}{r_{o2} + (R_1 \parallel R_2)} \tag{11.52}$$

Combining Eqs. (11.51) and (11.52) results in A:

$$A \equiv \frac{I_o}{I_i} = -\mu \frac{R_i}{1/g_m + (R_1 \parallel R_2 \parallel r_{o2})} \frac{r_{o2}}{r_{o2} + (R_1 \parallel R_2)} \tag{11.53}$$

which with some manipulation can be expressed in the form

$$A = -\mu \frac{R_i}{(R_1 \parallel R_2)} \frac{(R_1 \parallel R_2 \parallel r_{o2})}{1/g_m + (R_1 \parallel R_2 \parallel r_{o2})} \tag{11.54}$$

Noting that R_o is the output resistance of Q, which has a resistance $(R_1 \parallel R_2)$ in its source lead, we can write

$$R_o = r_{o2} + (R_1 \parallel R_2) + (g_m r_{o2})(R_1 \parallel R_2) \tag{11.55}$$

(c) The loop gain is obtained by combining Eqs. (11.48) and (11.53),

$$A\beta = \mu \frac{R_i}{R_2} \frac{(R_1 \parallel R_2 \parallel r_{o2})}{1/g_m + (R_1 \parallel R_2 \parallel r_{o2})} \tag{11.56}$$

The closed-loop gain A_f can be obtained by substituting the expressions of A and $A\beta$ (Eqs. 11.54 and 11.56) in $A_f = A/(1 + A\beta)$.

The input resistance R_{if} is found as

$$R_{if} = R_i/(1 + A\beta)$$

Since by definition,

$$R_{if} = R_s \parallel R_{\text{in}}$$

we can easily find R_{in}.

For the output resistance, we have

$$R_{\text{out}} = R_{of} = R_o(1 + A\beta)$$

(d) For the numerical values given,

$$R_i = \infty \parallel \infty \parallel (10 + 90) = 100 \text{ k}\Omega$$

Example 11.10 *continued*

The open-loop gain A can be determined by using Eq. (11.54),

$$A = -10,764 \text{ A/A}$$

and the feedback factor β can be found by using Eq. (11.48),

$$\beta = -\frac{R_1}{R_1 + R_2} = -\frac{10}{10+90} = -0.1 \text{ A/A}$$

Thus,

$$A\beta = 1076.4$$

and

$$A_f = -\frac{10,764}{1+1076.4} = -9.99 \text{ A/A}$$

which is very close to the ideal value of

$$A_f \simeq -\left(1+\frac{R_2}{R_1}\right) = -\left(1+\frac{90}{10}\right) = -10 \text{ A/A}$$

$$R_{if} = \frac{R_i}{1+A\beta} = \frac{100 \text{ k}\Omega}{1+1076.4} = 92.8 \ \Omega$$

since $R_s = \infty$,

$$R_{in} = R_{if} = 92.8 \ \Omega$$

$$R_o = r_{o2} + (R_1 \| R_2) + g_m r_{o2}(R_1 \| R_2) = 929 \text{ k}\Omega$$

$$R_{out} = R_{of} = (1+A\beta)R_o = 1077.4 \times 929 = 1001 \text{ M}\Omega$$

EXERCISES

11.20 For the amplifier in Example 11.10, find the values of A_f, R_{in}, and R_{out} when the value of μ is 10 times lower, that is, when $\mu = 100$.

Ans. −9.91 A/A; 920 Ω; 101 MΩ

11.21 If in the circuit in Fig. 11.27(a), R_2 is short-circuited, find the ideal value of A_f. For the case $R_s = R_{id} = \infty$, give expressions for R_i, R_o, A, β, A_f, R_{in}, and R_{out}.

Ans. $A_f = -1$ A/A; $R_i = R_1$; $R_o = r_{o2}$, $A = -\mu g_m R_1$; $\beta = -1$; $A_f = -\mu g_m R_1/(1+\mu g_m R_1)$; $R_{in} \simeq 1/\mu g_m$; $R_{out} \simeq \mu(g_m r_{o2})R_1$.

11.6 Summary of the Feedback-Analysis Method

Table 11.2 provides a summary of the rules and relationships employed in the analysis and design of the four types of feedback amplifier. In addition to the wealth of information in Table 11.2, we offer the following important analysis tips.

1. Always begin the analysis by determining an approximate value for the closed-loop gain A_f, assuming that the loop gain $A\beta$ is large and thus

$$A_f \simeq 1/\beta$$

   This value should serve as a check on the final value you find for A_f. How close the actual A_f is to this ideal value will depend on how large $A\beta$ is compared to unity.
2. The shunt connection at input or output always results in reducing the corresponding resistance (input or output). The series connection at input or output always results in increasing the corresponding resistance (input or output).
3. In utilizing negative feedback to improve the properties of an amplifier under design, the starting point in the design is the selection of the feedback topology appropriate for the application at hand. Then the required amount of negative feedback $(1+A\beta)$ can be ascertained utilizing the fact that it is this quantity that determines the magnitude of improvement in the various amplifier parameters. Also, the feedback factor β can be determined from the required closed-loop gain A_f,

$$\beta \simeq 1/A_f$$

11.7 The Stability Problem

11.7.1 Transfer Function of the Feedback Amplifier

In a feedback amplifier such as that represented by the general structure of Fig. 11.1, the open-loop gain A is generally a function of frequency, and it should therefore be more accurately called the **open-loop transfer function**, $A(s)$. Also, we have been assuming for the most part that the feedback network is resistive and hence that the feedback factor β is constant, but this need not be always the case. We shall therefore assume that in the general case the **feedback transfer function** is $\beta(s)$. It follows that the **closed-loop transfer function** $A_f(s)$ is given by

$$A_f(s) = \frac{A(s)}{1+A(s)\beta(s)} \tag{11.57}$$

To focus attention on the points central to our discussion in this section, we shall assume that the amplifier is direct coupled with constant dc gain A_0 and with poles and zeros occurring in the high-frequency band. Also, for the time being let us assume that at low frequencies $\beta(s)$ reduces to a constant value. Thus at low frequencies the loop gain $A(s)\beta(s)$ becomes a constant, which should be a positive number; otherwise the feedback would not be negative. The question then is: What happens at higher frequencies?

Table 11.2 Summary of Relationships for the Four Feedback-Amplifier Topologies

Feedback Amplifier	Feedback Topology	x_i	x_o	x_f	x_s	A	β	A_f	Source Form	Loading of Feedback Network Is Obtained: At Input	Loading of Feedback Network Is Obtained: At Output	To Find β, Apply to Port 2 of Feedback Network	R_{if}	R_{of}	Refer to Figs.
Voltage	Series–shunt	V_i	V_o	V_f	V_s	$\frac{V_o}{V_i}$	$\frac{V_f}{V_o}$	$\frac{V_o}{V_s}$	Thévenin	By short-circuiting port 2 of feedback network	By open-circuiting port 1 of feedback network	a voltage, and find the open-circuit voltage at port 1	$R_i(1+A\beta)$	$\frac{R_o}{1+A\beta}$	11.12 11.14
Current	Shunt–series	I_i	I_o	I_f	I_s	$\frac{I_o}{I_i}$	$\frac{I_f}{I_o}$	$\frac{I_o}{I_s}$	Norton	By open-circuiting port 2 of feedback network	By short-circuiting port 1 of feedback network	a current, and find the short-circuit current at port 1	$\frac{R_i}{1+A\beta}$	$R_o(1+A\beta)$	11.25 11.26
Transconductance	Series–series	V_i	I_o	V_f	V_s	$\frac{I_o}{V_i}$	$\frac{V_f}{I_o}$	$\frac{I_o}{V_s}$	Thévenin	By open-circuiting port 2 of feedback network	By open-circuiting port 1 of feedback network	a current, and find the open-circuit voltage at port 1	$R_i(1+A\beta)$	$R_o(1+A\beta)$	11.18 11.19
Transresistance	Shunt–shunt	I_i	V_o	I_f	I_s	$\frac{V_o}{I_i}$	$\frac{I_f}{V_o}$	$\frac{V_o}{I_s}$	Norton	By short-circuiting port 2 of feedback network	By short-circuiting port 1 of feedback network	a voltage, and find the short-circuit current at port 1	$\frac{R_i}{1+A\beta}$	$\frac{R_o}{1+A\beta}$	11.22 11.23

For physical frequencies $s = j\omega$, Eq. (11.57) becomes

$$A_f(j\omega) = \frac{A(j\omega)}{1 + A(j\omega)\beta(j\omega)} \tag{11.58}$$

Thus the loop gain $A(j\omega)\beta(j\omega)$ is a complex number that can be represented by its magnitude and phase,

$$\begin{aligned} L(j\omega) &\equiv A(j\omega)\beta(j\omega) \\ &= |A(j\omega)\beta(j\omega)|e^{j\phi(\omega)} \end{aligned} \tag{11.59}$$

It is the manner in which the loop gain varies with frequency that determines the stability or instability of the feedback amplifier. To appreciate this fact, consider the frequency at which the phase angle $\phi(\omega)$ becomes 180°. At this frequency, ω_{180}, the loop gain $A(j\omega)\beta(j\omega)$ will be a real number with a negative sign. Thus at this frequency the feedback will become positive. If at $\omega = \omega_{180}$ the magnitude of the loop gain is less than unity, then from Eq. (11.58) we see that the closed-loop gain $A_f(j\omega)$ will be greater than the open-loop gain $A(j\omega)$, since the denominator of Eq. (11.58) will be smaller than unity. Nevertheless, the feedback amplifier will be stable.

On the other hand, if at the frequency ω_{180} the magnitude of the loop gain is equal to unity, it follows from Eq. (11.58) that $A_f(j\omega)$ will be infinite. This means that the amplifier will have an output for zero input; this is by definition an **oscillator**. To visualize how this feedback loop may oscillate, consider the general loop of Fig. 11.1 with the external input x_s set to zero. Any disturbance in the circuit, such as the closure of the power-supply switch, will generate a signal $x_i(t)$ at the input to the amplifier. Such a noise signal usually contains a wide range of frequencies, and we shall now concentrate on the component with frequency $\omega = \omega_{180}$, that is, the signal $X_i \sin(\omega_{180}t)$. This input signal will result in a feedback signal given by

$$X_f = A(j\omega_{180})\beta(j\omega_{180})X_i = -X_i$$

Since X_f is further multiplied by –1 in the summer block at the input, we see that the feedback causes the signal X_i at the amplifier input to be *sustained*. That is, from this point on, there will be sinusoidal signals at the amplifier input and output of frequency ω_{180}. Thus the amplifier is said to oscillate at the frequency ω_{180}.

The question now is: What happens if at ω_{180} the magnitude of the loop gain is greater than unity? We shall answer this question, not in general, but for the restricted yet very important class of circuits in which we are interested here. The answer, which is not obvious from Eq. (11.58), is that the circuit will oscillate, and the oscillations will grow in amplitude until some nonlinearity (which is always present in some form) reduces the magnitude of the loop gain to exactly unity, at which point sustained oscillations will be obtained. This mechanism for starting oscillations by using positive feedback with a loop gain greater than unity, and then using a nonlinearity to reduce the loop gain to unity at the desired amplitude, will be exploited in the design of sinusoidal oscillators in Chapter 18. Our objective here is just the opposite: Now that we know how oscillations could occur in a negative-feedback amplifier, we wish to find methods to prevent their occurrence.

11.7.2 The Nyquist Plot

The Nyquist plot is a formalized approach for testing for stability based on the discussion above. It is simply a polar plot of loop gain, with frequency used as a parameter. Figure 11.28

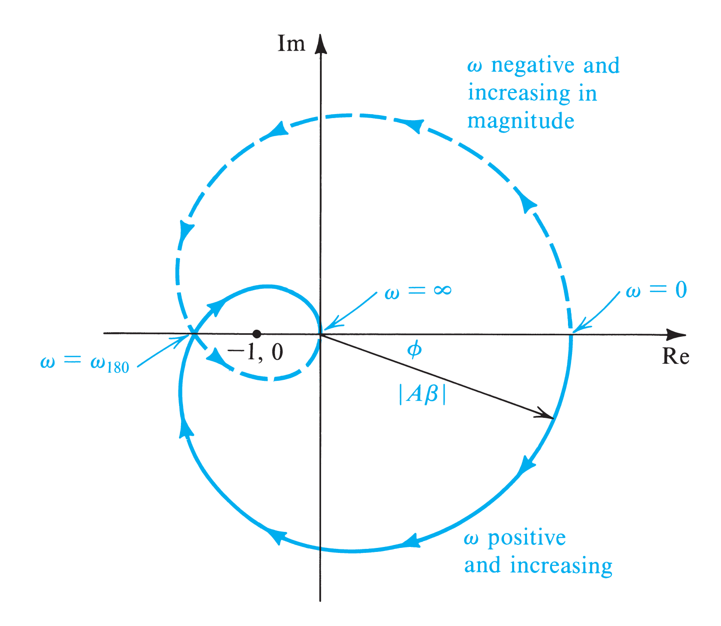


Figure 11.28 The Nyquist plot of an unstable amplifier.

shows such a plot. Note that the radial distance is $|A\beta|$ and the angle is the phase angle ϕ. The solid-line plot is for positive frequencies. Since the loop gain—and for that matter any gain function of a physical network—has a magnitude that is an even function of frequency and a phase that is an odd function of frequency, the $A\beta$ plot for negative frequencies (shown in Fig. 11.28 as a broken line) can be drawn as a mirror image through the Re axis.

The Nyquist plot intersects the negative real axis at the frequency ω_{180}. Thus, if this intersection occurs to the left of the point (–1, 0), we know that the magnitude of loop gain at this frequency is greater than unity and the amplifier will be unstable. On the other hand, if the intersection occurs to the right of the point (–1, 0) the amplifier will be stable. It follows that if the Nyquist plot *encircles* the point (–1, 0) then the amplifier will be unstable. It should be mentioned, however, that this statement is a simplified version of the **Nyquist criterion**; nevertheless, it applies to all the circuits in which we are interested. For the full theory behind the Nyquist method and for details of its application, consult Haykin (1970).

EXERCISE

11.22 Consider a feedback amplifier for which the open-loop transfer function $A(s)$ is given by

$$A(s) = \left(\frac{10}{1 + s/10^4}\right)^3$$

Let the feedback factor β be a constant independent of frequency. Find the frequency ω_{180} at which the phase shift is 180°. Then, show that the feedback amplifier will be stable if the feedback

factor β is less than a critical value β_{cr} and unstable if $\beta \geq \beta_{cr}$, and find the value of β_{cr}. Hence, find the minimum value of the closed-loop gain for which the amplifier is stable.
Ans. $\omega_{180} = \sqrt{3} \times 10^4$ rad/s; $\beta_{cr} = 0.008$; $A_{f\,min} = 111.1$

HARRY NYQUIST—A DIVERSE ELECTRONICS FUNDAMENTALIST:

Harry Nyquist, a Swedish-born electrical engineer working for Bell Labs and its predecessor, was responsible for developments in communications electronics involving thermal noise, feedback-amplifier stability, telegraphy, facsimile, television, and many other areas.

In *The Idea Factory*, an excellent book on the history of Bell Labs, Jon Gertner notes that

> [S]ome lawyers in the patent office of the Bell Labs decided to study whether there was an organizing principle that could explain why certain individuals were more productive than others. They discerned only one common thread: Workers with the most patents often shared lunch or breakfast with a Bell Labs electrical engineer named Harry Nyquist. It wasn't the case that Nyquist gave them specific ideas. Rather, as one scientist recalled, "he drew people out, got them thinking." More than anything, Nyquist asked good questions.

11.8 Effect of Feedback on the Amplifier Poles

The amplifier frequency response and stability are determined directly by its poles. Therefore we shall investigate the effect of feedback on the poles of the amplifier.[7]

11.8.1 Stability and Pole Location

We shall begin by considering the relationship between stability and pole location. For an amplifier or any other system to be stable, its poles should lie in the left half of the s plane. A pair of complex-conjugate poles on the $j\omega$ axis gives rise to sustained sinusoidal oscillations. Poles in the right half of the s plane give rise to growing oscillations.

To verify the statement above, consider an amplifier with a pole pair at $s = \sigma_0 \pm j\omega_n$. If this amplifier is subjected to a disturbance, such as that caused by closure of the power-supply switch, its transient response will contain terms of the form

$$v(t) = e^{\sigma_0 t}\left[e^{+j\omega_n t} + e^{-j\omega_n t}\right] = 2e^{\sigma_0 t}\cos(\omega_n t) \tag{11.60}$$

This is a sinusoidal signal with an envelope $e^{\sigma_0 t}$. Now if the poles are in the left half of the s plane, then σ_0 will be negative and the oscillations will decay exponentially toward zero, as shown in Fig. 11.29(a), indicating that the system is stable. If, on the other hand, the poles are in

[7]For a brief review of poles and zeros and related concepts, refer to Appendix F.

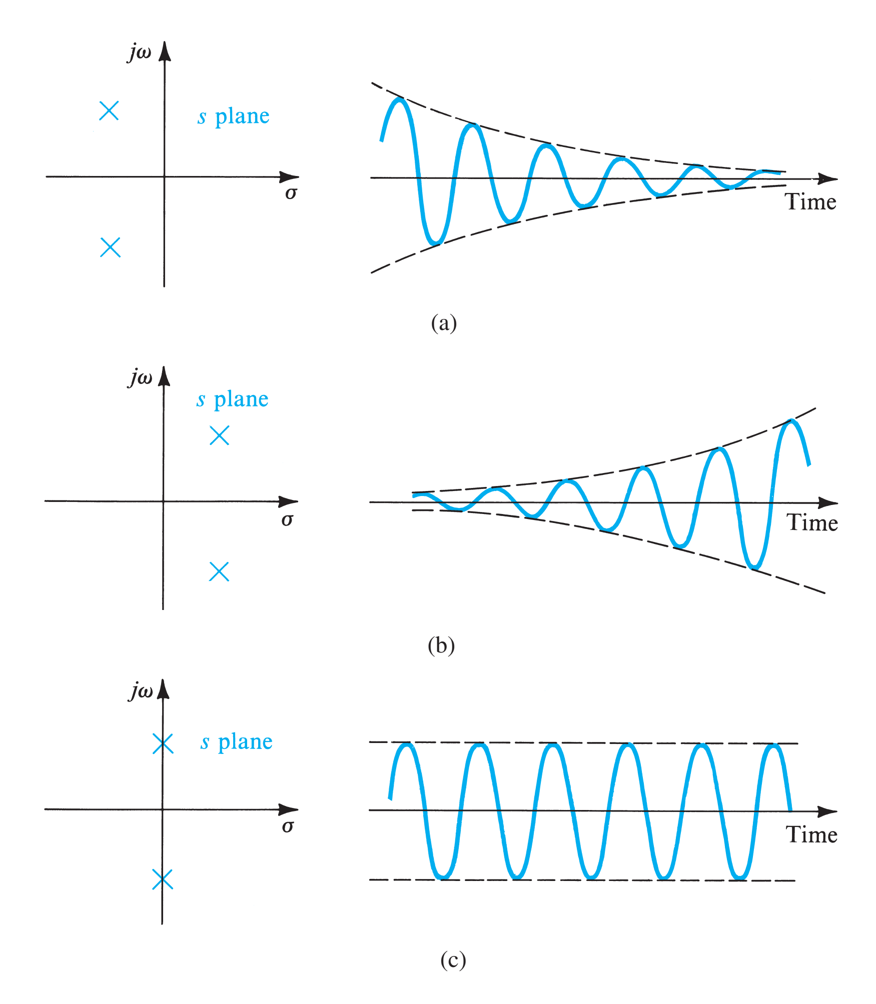


Figure 11.29 Relationship between pole location and transient response.

the right half-plane, then σ_0 will be positive, and the oscillations will grow exponentially (until some nonlinearity limits their growth), as shown in Fig. 11.29(b). Finally, if the poles are on the $j\omega$ axis, then σ_0 will be zero and the oscillations will be sustained, as shown in Fig. 11.29(c).

Although the discussion above is in terms of complex-conjugate poles, it can be shown that the existence of any right-half-plane poles results in instability.

11.8.2 Poles of the Feedback Amplifier

From the closed-loop transfer function in Eq. (11.57), we see that the poles of the feedback amplifier are the zeros of $1 + A(s)\beta(s)$. That is, the feedback amplifier poles are obtained by solving the equation



$$1 + A(s)\beta(s) = 0 \qquad (11.61)$$

which is called the **characteristic equation** of the feedback loop. It should therefore be apparent that applying feedback to an amplifier changes its poles.

In the following, we shall consider how feedback affects the amplifier poles. For this purpose we shall assume that the open-loop amplifier has real poles and no finite zeros (i.e., all the zeros are at $s = \infty$). This will simplify the analysis and enable us to focus our attention on the fundamental concepts involved. We shall also assume that the feedback factor β is independent of frequency.

11.8.3 Amplifier with a Single-Pole Response

Consider first the case of an amplifier whose open-loop transfer function is characterized by a single pole:

$$A(s) = \frac{A_0}{1 + s/\omega_P} \tag{11.62}$$

The closed-loop transfer function is given by

$$A_f(s) = \frac{A_0/(1 + A_0\beta)}{1 + s/\omega_P(1 + A_0\beta)} \tag{11.63}$$

Thus the feedback moves the pole along the negative real axis to a frequency ω_{Pf},

$$\omega_{Pf} = \omega_P(1 + A_0\beta) \tag{11.64}$$

This process is illustrated in Fig. 11.30(a). Figure 11.30(b) shows Bode plots for $|A|$ and $|A_f|$. Note that while at low frequencies the difference between the two plots is 20 log$(1 + A_0\beta)$, the two curves coincide at high frequencies. One can show that this indeed is the case by approximating Eq. (11.63) for frequencies $\omega \gg \omega_P(1 + A_0\beta)$:

$$A_f(s) \simeq \frac{A_0\omega_P}{s} \simeq A(s) \tag{11.65}$$

Physically speaking, at such high frequencies the loop gain is much smaller than unity and the feedback is ineffective.

Figure 11.30(b) clearly illustrates the fact that applying negative feedback to an amplifier results in extending its bandwidth at the expense of a reduction in gain. Since the pole of the

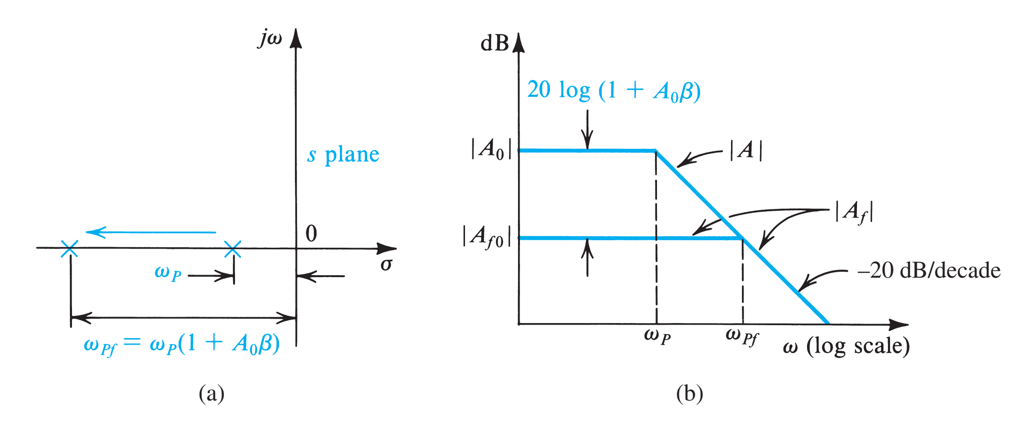


Figure 11.30 Effect of feedback on **(a)** the pole location and **(b)** the frequency response of an amplifier having a single-pole, open-loop response.

closed-loop amplifier never enters the right half of the s plane, the single-pole amplifier is stable for any value of β. Thus this amplifier is said to be **unconditionally stable**. This result, however, is hardly surprising, since the phase lag associated with a single-pole response can never be greater than 90°. Thus the loop gain never achieves the 180° phase shift required for the feedback to become positive.

EXERCISE

11.23 An op amp having a single-pole rolloff at 100 Hz and a low-frequency gain of 10^5 is operated in a feedback loop with $\beta = 0.01$. What is the factor by which feedback shifts the pole? To what frequency? If β is changed to a value that results in a nominal closed-loop gain of +1, to what frequency does the pole shift?

Ans. 1001; 100.1 kHz; 10 MHz

11.8.4 Amplifier with a Two-Pole Response

Consider next an amplifier whose open-loop transfer function is characterized by two real-axis poles:

$$A(s) = \frac{A_0}{(1 + s/\omega_{P1})(1 + s/\omega_{P2})} \tag{11.66}$$

In this case, the closed-loop poles are obtained from $1 + A(s)\beta = 0$, which leads to

$$s^2 + s(\omega_{P1} + \omega_{P2}) + (1 + A_0\beta)\omega_{P1}\omega_{P2} = 0 \tag{11.67}$$

Thus the closed-loop poles are given by

$$s = -\frac{1}{2}(\omega_{P1} + \omega_{P2}) \pm \frac{1}{2}\sqrt{(\omega_{P1} + \omega_{P2})^2 - 4(1 + A_0\beta)\omega_{P1}\omega_{P2}} \tag{11.68}$$

From Eq. (11.68) we see that as the loop gain $A_0\beta$ is increased from zero, the poles are brought closer together. Then a value of loop gain is reached at which the poles become coincident. If the loop gain is further increased, the poles become complex conjugate and move along a vertical line. Figure 11.31 shows the locus of the poles for increasing loop gain. This plot is called a **root-locus diagram**, where "root" refers to the fact that the poles are the roots of the characteristic equation.

From the root-locus diagram of Fig. 11.31 we see that this feedback amplifier also is unconditionally stable. Again, this result should come as no surprise; the maximum phase shift of $A(s)$ in this case is 180° (90° per pole), but this value is reached at $\omega = \infty$. Thus there is no finite frequency at which the phase shift reaches 180°.

Another observation to make on the root-locus diagram of Fig. 11.31 is that the open-loop amplifier might have a dominant pole, but this is not necessarily the case for the closed-loop amplifier. The response of the closed-loop amplifier can, of course, always be plotted once the poles have been found from Eq. (11.68). As is the case with second-order responses generally, the closed-loop response can show a peak (see Chapter 17). To be more specific,

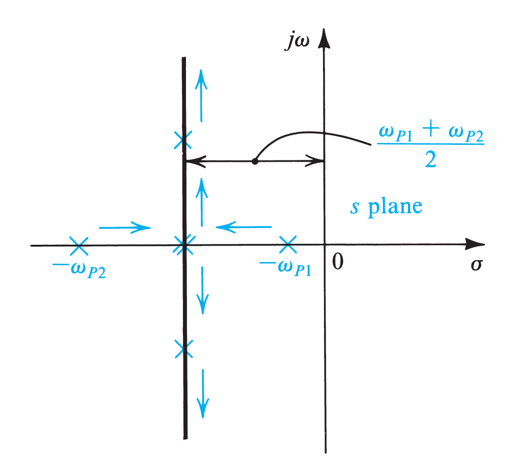


Figure 11.31 Root-locus diagram for a feedback amplifier whose open-loop transfer function has two real poles.

the characteristic equation of a second-order network can be written in the standard form

$$s^2 + s\frac{\omega_0}{Q} + \omega_0^2 = 0 \tag{11.69}$$

where ω_0 is called the **pole frequency** and Q is called **pole Q factor**. The poles are complex if Q is greater than 0.5. A geometric interpretation for ω_0 and Q of a pair of complex-conjugate poles is given in Fig. 11.32, from which we note that ω_0 is the radial distance of the poles from the origin and that Q indicates the distance of the poles from the $j\omega$ axis. Poles on the $j\omega$ axis have $Q = \infty$.

By comparing Eqs. (11.67) and (11.69), we obtain the Q factor for the poles of the feedback amplifier as

$$Q = \frac{\sqrt{(1 + A_0\beta)\omega_{P1}\omega_{P2}}}{\omega_{P1} + \omega_{P2}} \tag{11.70}$$

From the study of second-order network responses in Chapter 17, it will be seen that the response of the feedback amplifier under consideration shows no peaking for $Q \leq 0.707$. The boundary case corresponding to $Q = 0.707$ (poles at 45° angles) results in the **maximally flat** response. Figure 11.33 shows a number of possible responses obtained for various values of Q (or, correspondingly, various values of $A_0\beta$).

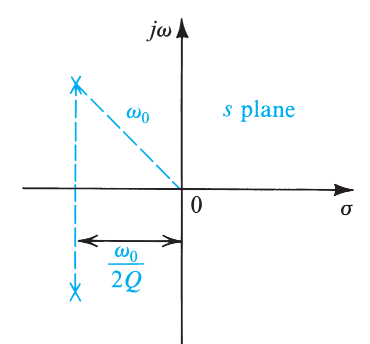


Figure 11.32 Definition of ω_0 and Q of a pair of complex-conjugate poles.

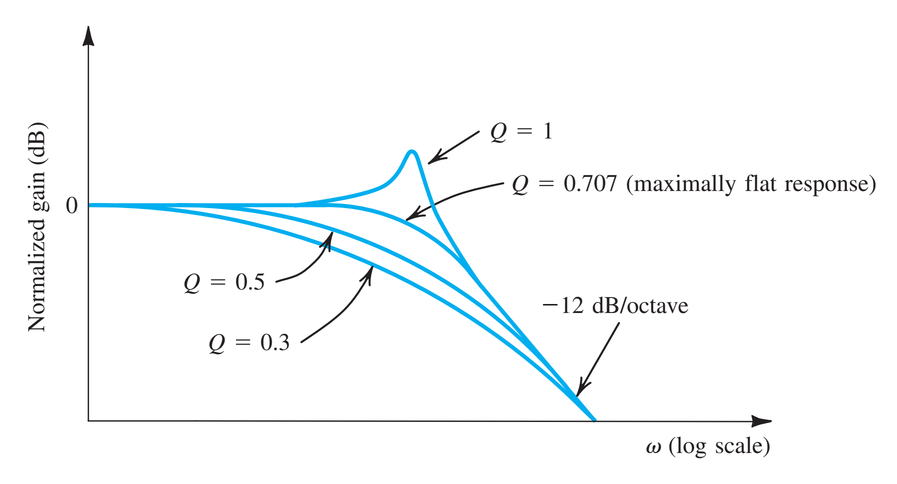


Figure 11.33 Normalized gain of a two-pole feedback amplifier for various values of Q. Note that Q is determined by the loop gain according to Eq. (11.70).

EXERCISE

11.24 An amplifier with a low-frequency gain of 100 and poles at 10^4 rad/s and 10^6 rad/s is incorporated in a negative-feedback loop with feedback factor β. For what value of β do the poles of the closed-loop amplifier coincide? What is the corresponding Q of the resulting second-order system? For what value of β is a maximally flat response achieved? What is the low-frequency closed-loop gain in the maximally flat case?
Ans. 0.245; 0.5; 0.5; 1.96 V/V

Example 11.11

As an illustration of some of the ideas just discussed, we consider the positive-feedback circuit shown in Fig. 11.34(a). Find the loop transmission $L(s)$ and the characteristic equation. Sketch a root-locus diagram for varying K, and find the value of K that results in a maximally flat response and the value of K that makes the circuit oscillate. Assume that the amplifier has frequency-independent gain, infinite input impedance, and zero output impedance.

Solution

To obtain the loop transmission, we short-circuit the signal source and break the loop at the amplifier input. We then apply a test voltage V_t and find the returned voltage V_r, as indicated in Fig. 11.34(b). The

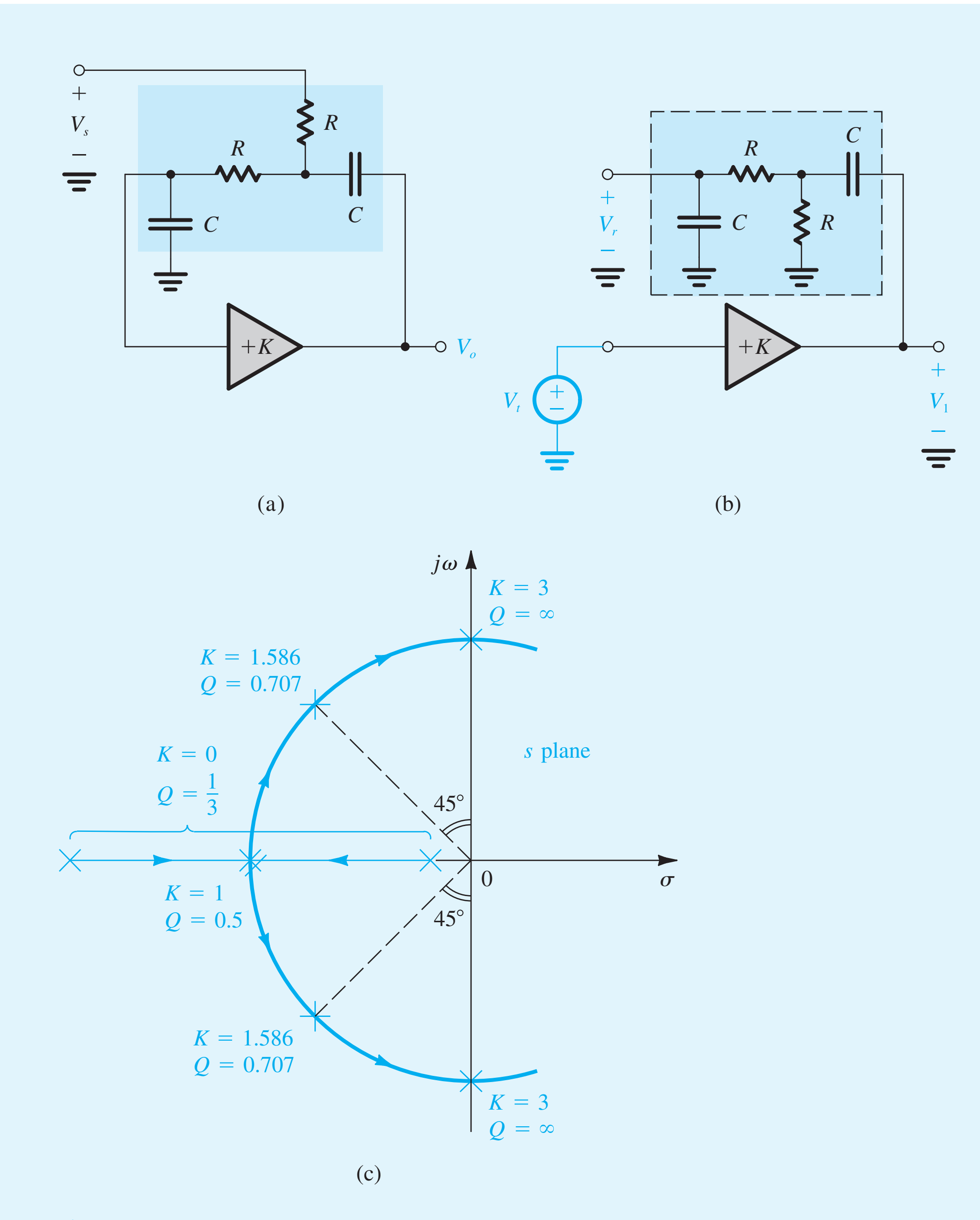


Figure 11.34 Circuits and plot for Example 11.11.

loop transmission $L(s) \equiv A(s)\beta(s)$ is given by

$$L(s) = -\frac{V_r}{V_t} = -KT(s) \tag{11.71}$$

Example 11.11 *continued*

where $T(s)$ is the transfer function of the two-port RC network shown inside the broken-line box in Fig. 11.34(b):

$$T(s) \equiv \frac{V_r}{V_1} = \frac{s(1/CR)}{s^2 + s(3/CR) + (1/CR)^2} \tag{11.72}$$

Thus,

$$L(s) = \frac{-s(K/CR)}{s^2 + s(3/CR) + (1/CR)^2} \tag{11.73}$$

The characteristic equation is

$$1 + L(s) = 0 \tag{11.74}$$

that is,

$$s^2 + s\frac{3}{CR} + \left(\frac{1}{CR}\right)^2 - s\frac{K}{CR} = 0$$

$$s^2 + s\frac{3-K}{CR} + \left(\frac{1}{CR}\right)^2 = 0 \tag{11.75}$$

By comparing this equation to the standard form of the second-order characteristic equation (Eq. 11.69), we see that the pole frequency ω_0 is given by

$$\omega_0 = \frac{1}{CR} \tag{11.76}$$

and the Q factor is

$$Q = \frac{1}{3-K} \tag{11.77}$$

Thus for $K = 0$, the poles have $Q = \frac{1}{3}$ and are therefore located on the negative real axis. As K is increased, the poles are brought closer together and eventually coincide ($Q = 0.5$, $K = 1$). Further increasing K results in the poles becoming complex and conjugate. The root locus is then a circle because the radial distance ω_0 remains constant (Eq. 11.76) independent of the value of K.

The maximally flat response is obtained when $Q = 0.707$, which results when $K = 1.586$. In this case the poles are at 45° angles, as indicated in Fig. 11.34(c). The poles cross the $j\omega$ axis into the right half of the s plane at the value of K that results in $Q = \infty$, that is, $K = 3$. Thus for $K \geq 3$ this circuit becomes unstable. This might appear to contradict our earlier conclusion that the feedback amplifier with a second-order response is unconditionally stable. Note, however, that the circuit in this example is quite different from the negative-feedback amplifier that we have been studying. Here we have an amplifier with a positive gain K and a feedback network whose transfer function $T(s)$ is frequency dependent. This feedback is in fact *positive,* and the circuit will oscillate at the frequency for which the phase of $T(j\omega)$ is zero (which is $1/CR$).

Example 11.11 illustrates the use of feedback (positive feedback in this case) to move the poles of an RC network from their negative real-axis locations to complex-conjugate locations. One can accomplish the same task using negative feedback, as the root-locus diagram of

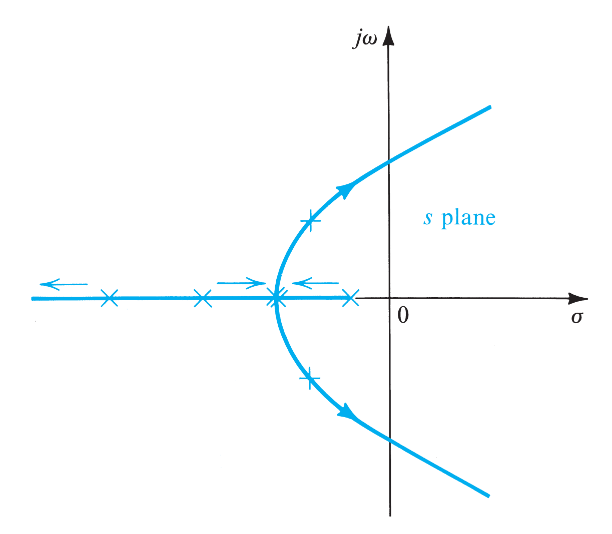


Figure 11.35 Root-locus diagram for an amplifier with three poles. The arrows indicate the pole movement as $A_0\beta$ is increased.

Fig. 11.31 demonstrates. The process of pole control is the essence of *active-filter design,* as will be discussed in Chapter 17.

11.8.5 Amplifiers with Three or More Poles

Figure 11.35 shows the root-locus diagram for a feedback amplifier whose open-loop response is characterized by three poles. As indicated, increasing the loop gain from zero moves the highest-frequency pole outward while the two other poles are brought closer together. As $A_0\beta$ is increased further, the two poles become coincident and then become complex and conjugate. A value of $A_0\beta$ exists at which this pair of complex-conjugate poles enters the right half of the s plane, thus causing the amplifier to become unstable.

This result is not entirely unexpected, since an amplifier with three poles has a phase shift that reaches –270° as ω approaches ∞. Thus there exists a finite frequency, ω_{180}, at which the loop gain has 180° phase shift.

From the root-locus diagram of Fig. 11.35, we observe that one can always maintain amplifier stability by keeping the loop gain $A_0\beta$ smaller than the value corresponding to the poles entering the right half-plane. In terms of the Nyquist diagram, the critical value of $A_0\beta$ is that for which the diagram passes through the (–1, 0) point. Reducing $A_0\beta$ below this value causes the Nyquist plot to shrink and thus intersect the negative real axis to the right of the (–1, 0) point, indicating stable amplifier performance. On the other hand, increasing $A_0\beta$ above the critical value causes the Nyquist plot to expand, thus encircling the (–1, 0) point and indicating unstable performance.

For a given open-loop gain A_0 the conclusions above can be stated in terms of the feedback factor β. That is, there exists a *maximum value* for β above which the feedback amplifier becomes unstable. Alternatively, we can state that there exists a *minimum value* for the closed-loop gain A_{f0} below which the amplifier becomes unstable. To obtain lower values of closed-loop gain, one needs therefore to alter the loop transfer function $L(s)$. This is the process known as *frequency compensation*. We shall study the theory and techniques of frequency compensation in Section 11.10.

Before leaving this section, we point out that construction of the root-locus diagram for amplifiers having three or more poles as well as finite zeros is an involved process for which a systematic procedure exists. However, such a procedure will not be presented here, and the interested reader should consult Haykin (1970). Although the root-locus diagram provides the amplifier designer with considerable insight, other, simpler techniques based on Bode plots can be effectively employed, as will be explained in Section 11.9.

EXERCISE

11.25 Consider a feedback amplifier for which the open-loop transfer function $A(s)$ is given by

$$A(s) = \left(\frac{10}{1 + s/10^4}\right)^3$$

Let the feedback factor β be frequency independent. Find the closed-loop poles as functions of β, and show that the root locus is that of Fig. E11.26. Also find the value of β at which the amplifier becomes unstable. (*Note:* This is the same amplifier that was considered in Exercise 11.22.)

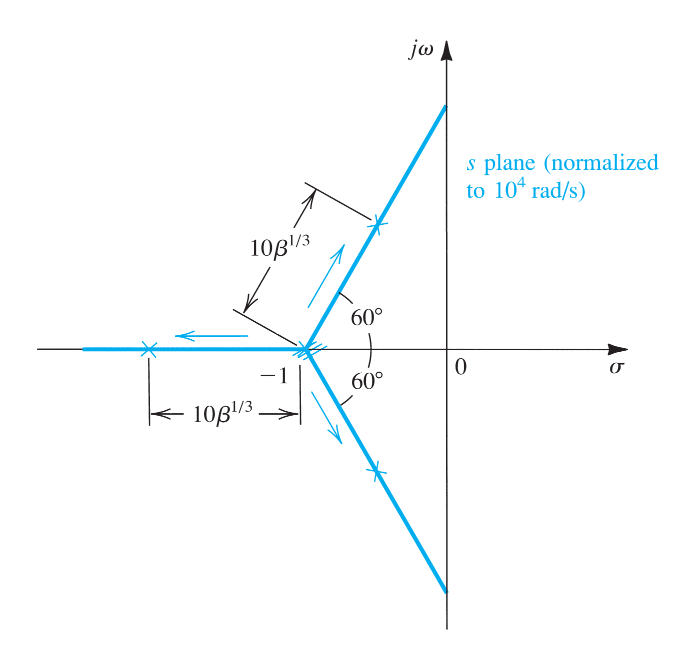


Figure E11.26

Ans. See Fig. E11.26; $\beta_{critical} = 0.008$

11.9 Stability Study Using Bode Plots

11.9.1 Gain and Phase Margins

From Sections 11.17 and 11.18 we know that whether a feedback amplifier is or is not stable can be determined by examining its loop gain $A\beta$ as a function of frequency. One of the simplest and most effective means for doing this is through the use of a Bode plot for $A\beta$, such as the one shown in Fig. 11.36. (Note that because the phase approaches –360°, the circuit examined is a fourth-order one.) The feedback amplifier whose loop gain is plotted in Fig. 11.36 will be stable, since at the frequency of 180° phase shift, ω_{180}, the magnitude of the loop gain is less than unity (negative dB). The difference between the value of $|A\beta|$ at ω_{180} and unity, called the **gain margin**, is usually expressed in decibels. The gain margin represents the amount by which the loop gain can be increased while stability is maintained. Feedback amplifiers are usually designed to have sufficient gain margin to allow for the inevitable changes in loop gain with temperature, time, and so on.

Another way to investigate the stability and to express its degree is to examine the Bode plot at the frequency for which $|A\beta| = 1$, which is the point at which the magnitude plot crosses the 0-dB line. If at this frequency the phase angle is less (in magnitude) than 180°, then the amplifier is stable. This is the situation illustrated in Fig. 11.36. The difference between the phase angle at this frequency and 180° is termed the **phase margin**. On the other hand, if at the frequency of unity loop-gain magnitude, the phase lag is in excess of 180°, the amplifier will be unstable.

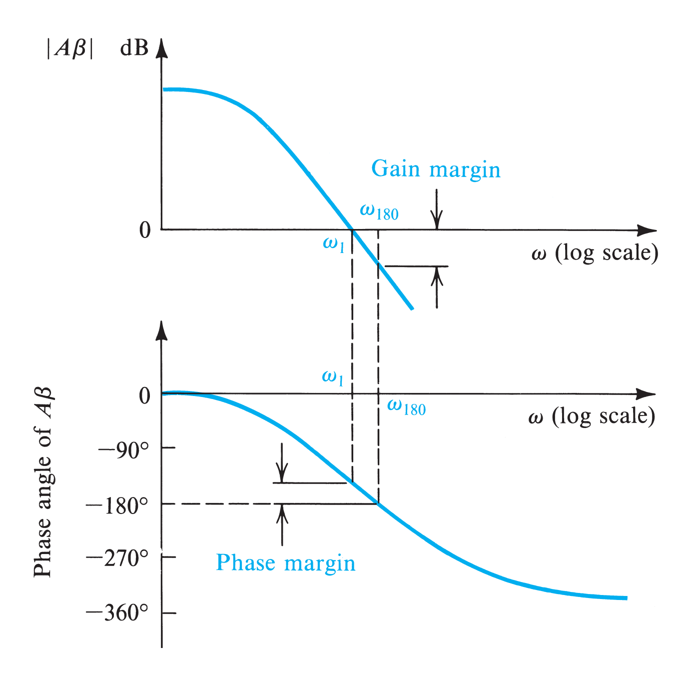


Figure 11.36 Bode plot for the loop gain $A\beta$ illustrating the definitions of the gain and phase margins.

EXERCISE

11.26 Consider an op amp having a single-pole, open-loop response with $A_0 = 10^5$ and $f_P = 10$ Hz. Let the op amp be ideal otherwise (infinite input impedance, zero output impedance, etc.). If this amplifier is connected in the noninverting configuration with a nominal low-frequency, closed-loop gain of 100, find the frequency at which $|A\beta| = 1$. Also, find the phase margin.
Ans. 10^4 Hz; 90°

11.9.2 Effect of Phase Margin on Closed-Loop Response

Feedback amplifiers are normally designed with a phase margin of at least 45°. The amount of phase margin has a profound effect on the shape of the closed-loop gain response. To see this relationship, consider a feedback amplifier with a large low-frequency loop gain, $A_0\beta \gg 1$. It follows that the closed-loop gain at low frequencies is approximately $1/\beta$. Denoting the frequency at which the magnitude of loop gain is unity by ω_1, we have (refer to Fig. 11.36)

$$A(j\omega_1)\beta = 1 \times e^{-j\theta} \tag{11.78}$$

where

$$\theta = 180° - \text{phase margin} \tag{11.79}$$

At ω_1 the closed-loop gain is

$$A_f(j\omega_1) = \frac{A(j\omega_1)}{1 + A(j\omega_1)\beta} \tag{11.80}$$

Substituting from Eq. (11.78) gives

$$A_f(j\omega_1) = \frac{(1/\beta)e^{-j\theta}}{1 + e^{-j\theta}} \tag{11.81}$$

Thus the magnitude of the gain at ω_1 is

$$|A_f(j\omega_1)| = \frac{1/\beta}{|1 + e^{-j\theta}|} \tag{11.82}$$

For a phase margin of 45°, $\theta = 135°$; and we obtain

$$|A_f(j\omega_1)| = 1.3\frac{1}{\beta} \tag{11.83}$$

That is, the gain peaks by a factor of 1.3 above the low-frequency value of $1/\beta$. This peaking increases as the phase margin is reduced, eventually reaching ∞ when the phase margin is zero. Zero phase margin, of course, implies that the amplifier can sustain oscillations [poles on the $j\omega$ axis; Nyquist plot passing through (–1, 0)].

EXERCISE

11.27 Find the closed-loop gain at ω_1 relative to the low-frequency gain when the phase margin is 30°, 60°, and 90°.
Ans. 1.93; 1; 0.707

11.9.3 An Alternative Approach for Investigating Stability

Investigating stability by constructing Bode plots for the loop gain $A\beta$ can be a tedious and time-consuming process, especially if we have to investigate the stability of a given amplifier for a variety of feedback networks. An alternative approach, which is much simpler, is to construct a Bode plot for the open-loop gain $A(j\omega)$ only. Assuming for the time being that β is independent of frequency, we can plot $20\log(1/\beta)$ as a horizontal straight line on the same plane used for $20\log|A|$. The difference between the two curves will be

$$20\log|A(j\omega)| - 20\log\frac{1}{\beta} = 20\log|A\beta| \tag{11.84}$$

which is the loop gain (in dB). We may therefore study stability by examining the difference between the two plots. If we wish to evaluate stability for a different feedback factor, we simply draw another horizontal straight line at the level $20\log(1/\beta)$.

To illustrate, consider an amplifier whose open-loop transfer function is characterized by three poles. For simplicity let the three poles be widely separated—say, at 0.1 MHz, 1 MHz, and 10 MHz, as shown in Fig. 11.37. Note that because the poles are widely separated, the phase is approximately –45° at the first pole frequency, –135° at the second, and –225° at the third. The frequency at which the phase of $A(j\omega)$ is –180° lies on the –40-dB/decade segment, as indicated in Fig. 11.37.

The open-loop gain of this amplifier can be expressed as

$$A = \frac{10^5}{(1+jf/10^5)(1+jf/10^6)(1+jf/10^7)} \tag{11.85}$$

from which $|A|$ can be easily determined for any frequency f (in Hz), and the phase can be obtained as

$$\phi = -\left[\tan^{-1}\left(f/10^5\right) + \tan^{-1}\left(f/10^6\right) + \tan^{-1}\left(f/10^7\right)\right] \tag{11.86}$$

The magnitude and phase graphs shown in Fig. 11.37 are obtained using the method for constructing Bode plots (Appendix F). These graphs provide approximate values for important amplifier parameters, with more exact values obtainable from Eqs. (11.85) and (11.86). For example, the frequency f_{180} at which the phase angle is 180° can be found from Fig. 11.37 to be approximately 3.2×10^6 Hz. Using this value as a starting point, a more exact value can be found by trial and error using Eq. (11.86). The result is $f_{180} = 3.34 \times 10^6$ Hz. At this frequency, Eq. (11.85) gives a gain magnitude of 58.2 dB, which is reasonably close to the approximate value of 60 dB given by Fig. 11.37.

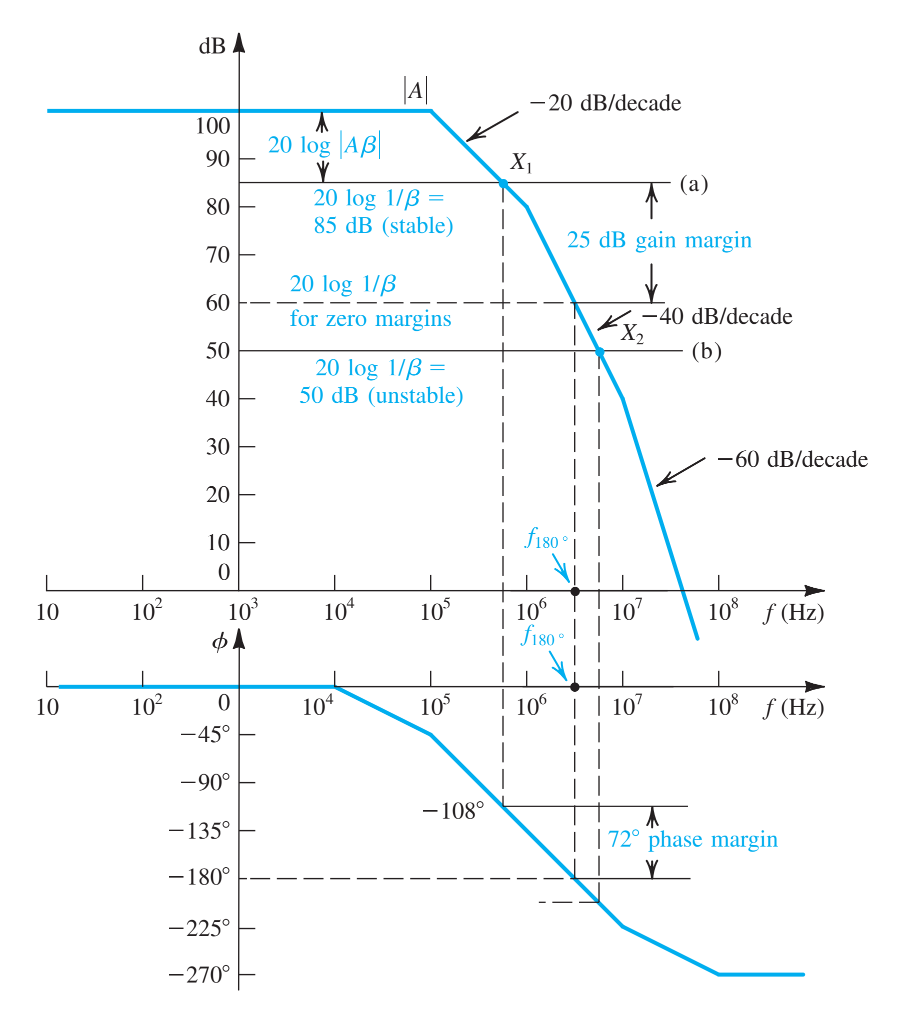


Figure 11.37 Stability analysis using Bode plot of $|A|$.

Consider next the straight line labeled (a) in Fig. 11.37. This line represents a feedback factor for which $20\log(1/\beta) = 85$ dB, which corresponds to $\beta = 5.623 \times 10^{-5}$ and a closed-loop gain of 83.6 dB. Since the loop gain is the difference between the $|A|$ curve and the $1/\beta$ line, the point of intersection X_1 corresponds to the frequency at which $|A\beta| = 1$. Using the graphs of Fig. 11.37, this frequency can be found to be approximately 5.6×10^5 Hz. A more exact value of 4.936×10^5 can be obtained using the transfer function equations. At this frequency the phase angle is approximately –108°. Thus the closed-loop amplifier, for which $20\log(1/\beta)$ = 85 dB, will be stable with a phase margin of 72°. The gain margin can be easily obtained from Fig. 11.37; it is 25 dB.

Next, suppose that we wish to use this amplifier to obtain a closed-loop gain of 50-dB nominal value. Since $A_0 = 100$ dB, we see that $A_0\beta \gg 1$ and $20\log(A_0\beta) \simeq 50$ dB, resulting in $20\log(1/\beta) \simeq 50$ dB. To see whether this closed-loop amplifier is or is not stable, we draw line (b) in Fig. 11.37 with a height of 50 dB. This line intersects the open-loop gain curve at point X_2, where the corresponding phase is greater than 180°. Thus the closed-loop amplifier with 50-dB gain will be unstable.

In fact, it can easily be seen from Fig. 11.37 that the *minimum* value of 20 log$(1/\beta)$ that can be used, with the resulting amplifier being stable, is 60 dB. In other words, the minimum value of stable closed-loop gain obtained with this amplifier is approximately 60 dB. At this value of gain, however, a manufactured version of this amplifier may still oscillate, since no margin is left to allow for possible changes in gain.

Since the 180°-phase point always occurs on the –40-dB/decade segment of the Bode plot for $|A|$, a rule of thumb to guarantee stability is as follows: *The closed-loop amplifier will be stable if the* 20 log$(1/\beta)$ *line intersects the* 20 log$|A|$ *curve at a point on the* –20-dB/decade *segment*. Following this rule ensures that a phase margin of at least 45° is obtained. For the example of Fig. 11.37, the rule implies that the maximum value of β is 10^{-4}, which corresponds to a closed-loop gain of approximately 80 dB.

The rule of thumb above can be generalized for the case in which β is a function of frequency. The general rule states that *at the intersection of* 20 log$[1/|\beta(j\omega)|]$ *and* 20 log$|A(j\omega)|$ *the difference of slopes* (called the **rate of closure**) *should not exceed* 20 dB/decade.

EXERCISE

11.28 Consider an op amp whose open-loop gain is identical to that of Fig. 11.37. Assume that the op amp is ideal otherwise. Let the op amp be connected as a differentiator. Use the rule of thumb above to show that for stable performance the differentiator time constant should be greater than 159 ms. [*Hint:* Recall that for a differentiator, the Bode plot for $1/|\beta(j\omega)|$ has a slope of +20 dB/decade and intersects the 0-dB line at $1/\tau$, where τ is the differentiator time constant.]

11.10 Frequency Compensation

In this section, we shall discuss methods for modifying the open-loop transfer function $A(s)$ of an amplifier having three or more poles so that the closed-loop amplifier is stable for a given desired value of closed-loop gain. This process is referred to as frequency compensation.

11.10.1 Theory

The simplest method of frequency compensation consists of introducing a new pole in the function $A(s)$ at a frequency, f_D, sufficiently low that the modified open-loop gain, $A'(s)$, intersects the 20 log$(1/|\beta|)$ curve with a slope difference of 20 dB/decade. As an example, let it be required to compensate the amplifier whose $A(s)$ is shown in Fig. 11.38 such that closed-loop amplifiers with β as high as 10^{-2} (i.e., closed-loop gains as low as approximately 40 dB) will be stable. First, we draw a horizontal straight line at the 40-dB level to represent 20 log$(1/\beta)$, as shown in Fig. 11.38. We then locate point Y on this line at the frequency of the first pole, f_{P1}. From Y we draw a line with –20-dB/decade slope and determine the point at which this line intersects the dc gain line, point Y'. This latter point

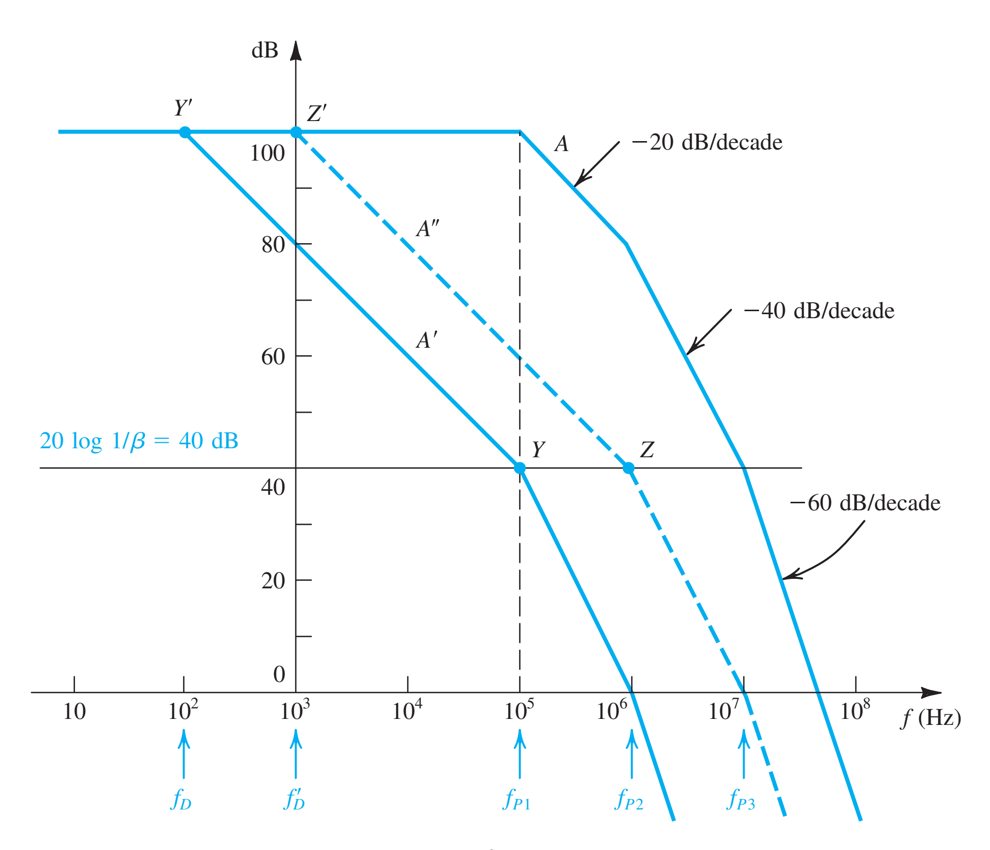


Figure 11.38 Frequency compensation for $\beta = 10^{-2}$. The response labeled A' is obtained by introducing an additional pole at f_D. The A'' response is obtained by moving the original low-frequency pole to f_D'.

gives the frequency f_D of the new pole that has to be introduced in the open-loop transfer function.

The compensated open-loop response $A'(s)$ is indicated in Fig. 11.38. It has four poles: at f_D, f_{P1}, f_{P2}, and f_{P3}. Thus $|A'|$ begins to roll off with a slope of –20 dB/decade at f_D. At f_{P1} the slope changes to –40 dB/decade, at f_{P2} it changes to –60 dB/decade, and so on. Since the 20 log$(1/\beta)$ line intersects the $20\log|A'|$ curve at point Y on the –20-dB/decade segment, the closed-loop amplifier with this β value (or lower values) will be stable.

A serious disadvantage of this compensation method is that at most frequencies the open-loop gain has been drastically reduced. This means that at most frequencies the amount of feedback available will be small. Since all the advantages of negative feedback are directly proportional to the amount of feedback, the performance of the compensated amplifier will be impaired.

Careful examination of Fig. 11.38 shows that the gain $A'(s)$ is low because of the pole at f_{P1}. If we can somehow eliminate this pole, then—rather than locating point Y, drawing YY', and so on—we can start from point Z (at the frequency of the second pole) and draw the line ZZ'. This would result in the open-loop curve $A''(s)$, which shows considerably higher gain than $A'(s)$.

Although it is not possible to eliminate the pole at f_{P1}, it is usually possible to shift that pole from $f = f_{P1}$ to $f = f_D'$. This makes the pole dominant and eliminates the need for introducing an additional lower-frequency pole, as will be explained next.

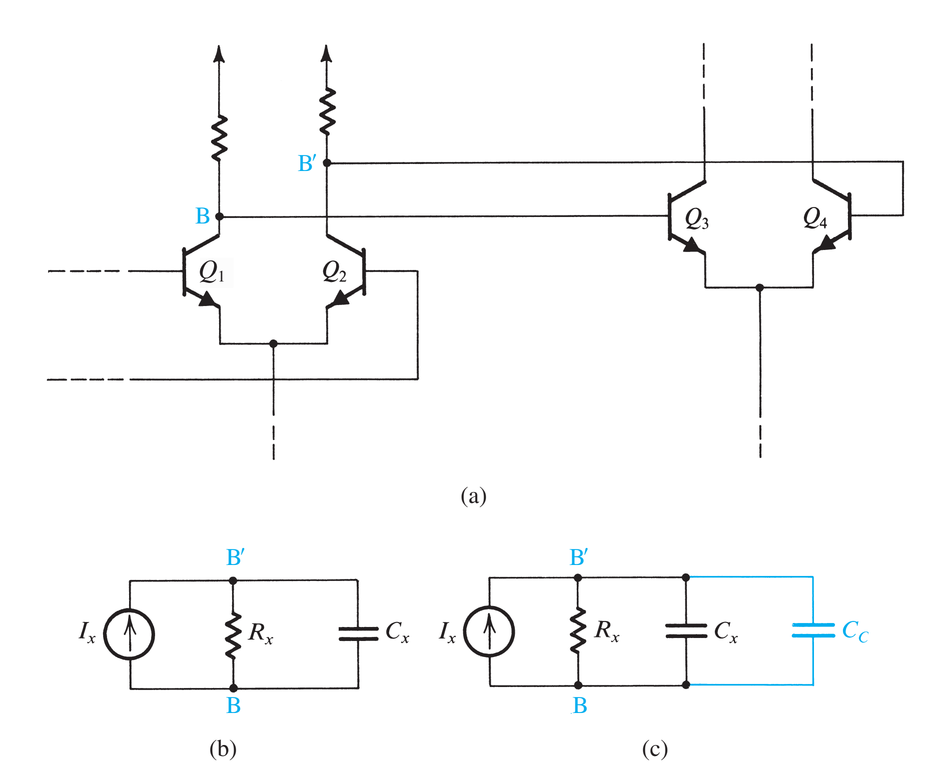


Figure 11.39 **(a)** Two cascaded gain stages of a multistage amplifier. **(b)** Equivalent circuit for the interface between the two stages in **(a)**. **(c)** Same circuit as in **(b)**, but with a compensating capacitor C_C added. Note that the analysis here applies equally well to MOS amplifiers.

11.10.2 Implementation

We shall now address the question of implementing the frequency-compensation scheme discussed above. The amplifier circuit normally consists of a number of cascaded gain stages, with each stage responsible for one or more of the transfer function poles. Through manual and/or computer analysis of the circuit, one identifies the stage that introduces each of the important poles f_{P1}, f_{P2}, and so on. For the purpose of our discussion, assume that the first pole f_{P1} is introduced at the interface between the two cascaded differential stages shown in Fig. 11.39(a). In Fig. 11.39(b) we show a simple small-signal model of the circuit at this interface. Current source I_x represents the output-signal current of the Q_1–Q_2 stage. Resistance R_x and capacitance C_x represent the total resistance and capacitance between the two nodes B and B′. It follows that the pole f_{P1} is given by

$$f_{P1} = \frac{1}{2\pi C_x R_x} \tag{11.87}$$

Let us now connect the compensating capacitor C_C between nodes B and B′. This will result in the modified equivalent circuit shown in Fig. 11.39(c), from which we see that the pole introduced will no longer be at f_{P1}; rather, the pole can be at any desired lower frequency f'_D :

$$f'_D = \frac{1}{2\pi (C_x + C_C) R_x} \tag{11.88}$$

We thus conclude that one can select an appropriate value for C_C to shift the pole frequency from f_{P1} to the value f'_D determined by point Z' in Fig. 11.38.

At this juncture it should be pointed out that adding the capacitor C_C will usually result in changes in the location of the other poles (those at f_{P2} and f_{P3}). One might therefore need to calculate the new location of f_{P2} and perform a few iterations to arrive at the required value for C_C.

A disadvantage of this implementation method is that the required value of C_C is usually quite large. Thus if the amplifier to be compensated is an IC op amp, it will be difficult, and probably impossible, to include this compensating capacitor on the IC chip. (As pointed out in Chapter 8 and in Appendix A, the maximum practical size of a monolithic capacitor is about 100 pF.) An elegant solution to this problem is to connect the compensating capacitor in the feedback path of an inverting amplifier stage. Because of the Miller effect (Section 10.3), the compensating capacitance will be multiplied by the stage gain, resulting in a much larger effective capacitance. Furthermore, as explained later, another unexpected benefit accrues.

11.10.3 Miller Compensation and Pole Splitting

Figure 11.40(a) shows one gain stage in a multistage amplifier. For simplicity, the stage is shown as a common-emitter amplifier, but in practice it can be a more elaborate circuit. In the feedback path of this common-emitter stage we have placed a compensating capacitor C_f.

Figure 11.40(b) shows a simplified equivalent circuit of the gain stage of Fig. 11.40(a). Here R_1 and C_1 represent the total resistance and total capacitance between node B and ground. Similarly, R_2 and C_2 represent the total resistance and total capacitance between node C and ground. Furthermore, it is assumed that C_1 includes the Miller component due to capacitance C_μ, and C_2 includes the input capacitance of the succeeding amplifier stage. Finally, I_i represents the output signal current of the preceding stage.

In the absence of the compensating capacitor C_f, we can see from Fig. 11.40(b) that there are two poles—one at the input and one at the output. Let us assume that these two poles are f_{P1} and f_{P2} of Fig. 11.38; thus,

$$f_{P1} = \frac{1}{2\pi C_1 R_1} \qquad f_{P2} = \frac{1}{2\pi C_2 R_2} \tag{11.89}$$

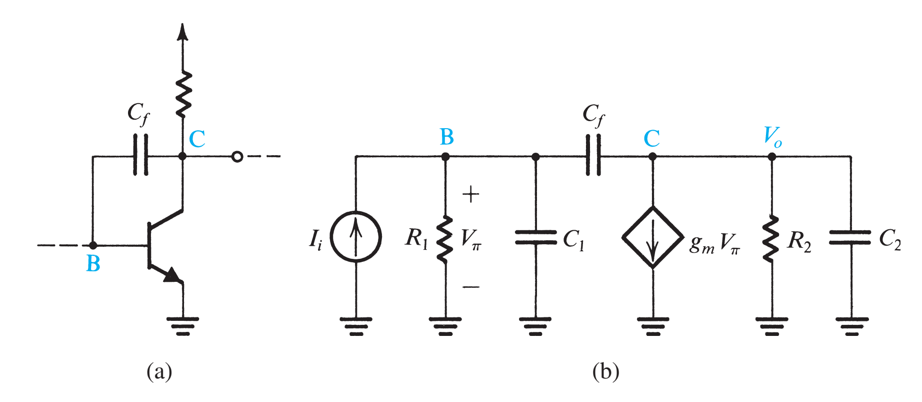


Figure 11.40 **(a)** A gain stage in a multistage amplifier with a compensating capacitor connected in the feedback path, and **(b)** equivalent circuit. Note that although a BJT is shown, the analysis applies equally well to the MOSFET case.

With C_f present, analysis of the circuit yields the transfer function

$$\frac{V_o}{I_i} = \frac{(sC_f - g_m)R_1R_2}{1 + s[C_1R_1 + C_2R_2 + C_f(g_mR_1R_2 + R_1 + R_2)] + s^2[C_1C_2 + C_f(C_1 + C_2)]R_1R_2} \quad (11.90)$$

The zero is usually at a much higher frequency than the dominant pole, and we shall neglect its effect. The denominator polynomial $D(s)$ can be written in the form

$$D(s) = \left(1 + \frac{s}{\omega'_{P1}}\right)\left(1 + \frac{s}{\omega'_{P2}}\right) = 1 + s\left(\frac{1}{\omega'_{P1}} + \frac{1}{\omega'_{P2}}\right) + \frac{s^2}{\omega'_{P1}\omega'_{P2}} \quad (11.91)$$

where ω'_{P1} and ω'_{P2} are the new frequencies of the two poles. Normally one of the poles will be dominant; $\omega'_{P1} \ll \omega'_{P2}$. Thus,

$$D(s) \simeq 1 + \frac{s}{\omega'_{P1}} + \frac{s^2}{\omega'_{P1}\omega'_{P2}} \quad (11.92)$$

Equating the coefficients of s in the denominator of Eq. (11.90) and in Eq. (11.92) results in

$$\omega'_{P1} = \frac{1}{C_1R_1 + C_2R_2 + C_f(g_mR_1R_2 + R_1 + R_2)}$$

which can be approximated by

$$\omega'_{P1} \simeq \frac{1}{g_mR_2C_fR_1} \quad (11.93)$$

To obtain ω'_{P2} we equate the coefficients of s^2 in the denominator of Eq. (11.90) and in Eq. (11.92) and use Eq. (11.93):

$$\omega'_{P2} \simeq \frac{g_mC_f}{C_1C_2 + C_f(C_1 + C_2)} \quad (11.94)$$

From Eqs. (11.93) and (11.94), we see that as C_f is increased, ω'_{P1} is reduced and ω'_{P2} is increased. This action is referred to as **pole splitting**. Note that the increase in ω'_{P2} is highly beneficial; it allows us to move point Z (see Fig. 11.38) further to the right, thus resulting in higher compensated open-loop gain. Finally, note from Eq. (11.93) that C_f is multiplied by the Miller effect factor g_mR_2, thus resulting in a much larger effective capacitance, $g_mR_2C_f$. In other words, the required value of C_f will be much smaller than that of C_C in Fig. 11.39.

Example 11.12

Consider an op amp whose open-loop transfer function is identical to that shown in Fig. 11.37. We wish to compensate this op amp so that the closed-loop amplifier with resistive feedback is stable for any gain (i.e., for β up to unity). Assume that the op-amp circuit includes a stage such as that of Fig. 11.40 with $C_1 = 100$ pF, $C_2 = 5$ pF, and $g_m = 40$ mA/V, that the pole at f_{P1} is caused by the input circuit of that stage, and that the pole at f_{P2} is introduced by the output circuit. Find the value of the compensating capacitor

Example 11.12 *continued*

for two cases: either if it is connected between the input node B and ground or in the feedback path of the transistor.

Solution

First we determine R_1 and R_2 from

$$f_{P1} = 0.1\text{ MHz} = \frac{1}{2\pi C_1 R_1}$$

Thus,

$$R_1 = \frac{10^5}{2\pi}\ \Omega$$

$$f_{P2} = 1\text{ MHz} = \frac{1}{2\pi C_2 R_2}$$

Thus,

$$R_2 = \frac{10^5}{\pi}\ \Omega$$

If a compensating capacitor C_C is connected across the input terminals of the transistor stage, then the frequency of the first pole changes from f_{P1} to f'_D:

$$f'_D = \frac{1}{2\pi\left(C_1 + C_C\right)R_1}$$

The second pole remains unchanged at 1 MHz. The required value for f'_D is determined by drawing a –20-dB/decade line from the 1-MHz frequency point on the $20\log(1/\beta) = 20\log 1 = 0$ dB line. This line will intersect the 100-dB dc gain line at 10 Hz. Thus,

$$f'_D = 10\text{ Hz} = \frac{1}{2\pi(C_1 + C_C)R_1}$$

which results in $C_C \simeq 1\ \mu\text{F}$, which is quite large and certainly cannot be included on the IC chip.

Next, if a compensating capacitor C_f is connected in the feedback path of the transistor, then both poles change location to the values given by Eqs. (11.93) and (11.94):

$$f'_{P1} \simeq \frac{1}{2\pi g_m R_2 C_f R_1} \qquad f'_{P2} \simeq \frac{g_m C_f}{2\pi\left[C_1 C_2 + C_f\left(C_1 + C_2\right)\right]} \tag{11.95}$$

To determine where we should locate the first pole, we need to know the value of f'_{P2}. As an approximation, let us assume that $C_f \gg C_2$, which enables us to obtain

$$f'_{P2} \simeq \frac{g_m}{2\pi\left(C_1 + C_2\right)} = 60.6\text{ MHz}$$

Thus it appears that this pole will move to a frequency higher than f_{P3} (which is 10 MHz). Let us therefore assume that the second pole will be at f_{P3}. This requires that the modified first pole be located at

$$f'_{P1} = \frac{f_{P3}}{A_0} = \frac{10^7\text{ Hz}}{10^5} = 100\text{ Hz}$$

Thus,

$$f'_{P1} = 100\,\text{Hz} = \frac{1}{2\pi g_m R_2 C_f R_1}$$

which results in $C_f = 78.5$ pF. Although this value is indeed much greater than C_2, we can determine the location of the pole f'_{P2} from Eq. (11.95), which yields $f'_{P2} = 57.2$ MHz, confirming that this pole has indeed been moved past f_{P3}.

We conclude that using Miller compensation not only results in a much smaller compensating capacitor but, owing to pole splitting, also enables us to place the dominant pole a decade higher in frequency. This results in a wider bandwidth for the compensated op amp.

EXERCISES

11.29 A multipole amplifier having a first pole at 1 MHz and an open-loop gain of 100 dB is to be compensated for closed-loop gains as low as 20 dB by the introduction of a new dominant pole. At what frequency must the new pole be placed?
Ans. 100 Hz

11.30 For the amplifier described in Exercise 11.29, rather than introducing a new dominant pole, we can use additional capacitance at the circuit node at which the first pole is formed to reduce the frequency of the first pole. If the frequency of the second pole is 10 MHz and if it remains unchanged while additional capacitance is introduced as mentioned, find the frequency to which the first pole must be lowered so that the resulting amplifier is stable for closed-loop gains as low as 20 dB. By what factor must the capacitance at the controlling node be increased?
Ans. 1000 Hz; 1000

Summary

- Negative feedback is employed to make the amplifier gain less sensitive to component variations; to control input and output resistances; to extend bandwidth; to reduce nonlinear distortion; and to enhance signal-to-interference ratio.
- The advantages above are obtained at the expense of a reduction in gain and at the risk of the amplifier becoming unstable (that is, oscillating). The latter problem is solved by careful design.
- The structure of an ideal negative-feedback amplifier is shown in Fig. 11.1. Table 11.1 summarizes the parameters and relationships governing the operation of the ideal structure.
- For each of the four basic types of amplifier, there is an appropriate feedback topology. The four topologies, together with their analysis procedure and their effects on input and output impedances, are summarized in Table 11.2 in Section 11.6.
- The key feedback parameters are the loop gain ($A\beta$), which for negative feedback must be a positive dimensionless number, and the amount of feedback ($1 + A\beta$). The latter directly determines gain reduction, gain

desensitivity, bandwidth extension, and changes in R_i and R_o.

- The loop gain $A\beta$ can be determined by breaking the feedback loop, as illustrated in Figs. 11.2 and 11.9. The value of $A\beta$ can be used together with the feedback factor β to determine A and hence A_f. This method, though simple, is incomplete as it does not enable the determination of the input and output resistances. For these, we utilize the systematic method for feedback analysis (refer to Table 11.2).
- The ideal or upper-bound value of the closed-loop gain A_f is $1/\beta$ and is approached when $A\beta \gg 1$.
- Since A and β are in general frequency dependent, the poles of the feedback amplifier are obtained by solving the characteristic equation $1 + A(s)\beta(s) = 0$.
- For the feedback amplifier to be stable, its poles must all be in the left half of the s plane.
- Stability is guaranteed if at the frequency for which the phase angle of $A\beta$ is 180° (i.e., ω_{180}), $|A\beta|$ is less than unity; the amount by which it is less than unity, expressed in decibels, is the gain margin. Alternatively, the amplifier is stable if, at the frequency at which $|A\beta| = 1$, the phase angle is less than 180°; the difference is the phase margin.
- The stability of a feedback amplifier can be analyzed by constructing a Bode plot for $|A|$ and superimposing on it a plot for $20 \log 1/|\beta|$. Stability is guaranteed if the two plots intersect with a difference in slope no greater than 6 dB/octave.
- To make a given amplifier stable for a given feedback factor β, the open-loop frequency response is suitably modified by a process known as frequency compensation.
- A popular method for frequency compensation involves connecting a feedback capacitor across an inverting stage in the amplifier. This causes the pole formed at the input of the amplifier stage to shift to a lower frequency and thus become dominant, while the pole formed at the output of the amplifier stage is moved to a very high frequency and thus becomes unimportant. This process is known as pole splitting.

PROBLEMS

Computer Simulation Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as allowable signal swing and amplifier nonlinear distortion. Instructions to assist in setting up PSpice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 11.1: The General Feedback Structure

11.1 A negative-feedback amplifier has a closed-loop gain $A_f = 200$ and an open-loop gain $A = 10^4$. What is the feedback factor β? If a manufacturing error results in a reduction of A to 10^3, what closed-loop gain results? What is the percentage change in A_f corresponding to this factor of 10 reduction in A?

11.2 Consider the op-amp circuit shown in Fig. P11.2, where the op amp has infinite input resistance and zero output resistance but finite open-loop gain A.

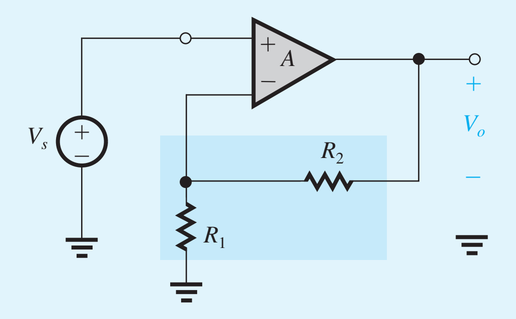


Figure P11.2

(a) Convince yourself that $\beta = R_1/(R_1 + R_2)$.

(b) If $R_1 = 10\ \text{k}\Omega$, find R_2 that results in $A_f = 10$ V/V for the following three cases: (i) $A = 1000$ V/V; (ii) $A = 200$ V/V; (iii) $A = 15$ V/V.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

(c) For each of the three cases in (b), find the percentage change in A_f that results when A decreases by 20%. Comment on the results.

11.3 The noninverting buffer op-amp configuration shown in Fig. P11.3 provides a direct implementation of the feedback loop of Fig. 11.1. Assuming that the op amp has infinite input resistance and zero output resistance, what is β? If $A = 1000$, what is the closed-loop voltage gain? What is the amount of feedback (in dB)? For $V_s = 1$ V, find V_o and V_i. If A decreases by 10%, what is the corresponding percentage decrease in A_f?

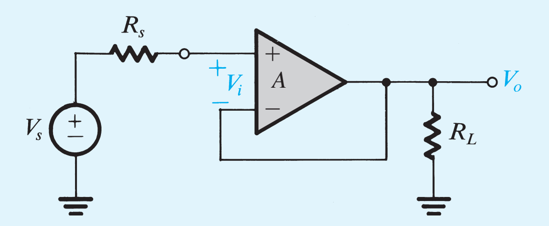


Figure P11.3

11.4 In a particular circuit represented by the block diagram of Fig. 11.1, a signal of 1 V from the source results in a difference signal of 10 mV being provided to the amplifying element A, and 5 V appearing at the output. For this arrangement, identify the values of A and β that apply.

11.5 (a) Show that in a negative-feedback amplifier with loop gain $A\beta \gg 1$, the closed-loop gain A_f is lower than its ideal value of $1/\beta$ by $(100/A\beta)$%.

(b) What is the minimum loop gain required so that A_f is within (i) 0.1%, (ii) 1%, and (iii) 5% of its ideal value?

11.6 In a particular amplifier design, the β network consists of a linear potentiometer for which β is 0.00 at one end, 1.00 at the other end, and 0.50 in the middle. As the potentiometer is adjusted, find the three values of closed-loop gain that result when the amplifier open-loop gain is (a) 1 V/V, (b) 10 V/V, (c) 100 V/V, (d) 1000 V/V, and (e) 10,000 V/V. Provide your results in a table in which there is a row for each value of A and a column for each value of β.

11.7 A newly constructed feedback amplifier undergoes a performance test with the following results: With the feedback connection removed, a source signal of 2 mV is required to provide a 5-V output; with the feedback connected, a 5-V output requires a 100-mV source signal. For this amplifier, identify values of A, β, $A\beta$, the closed-loop gain, and the amount of feedback (in dB).

D 11.8 An amplifier has an open-loop gain with a nominal value of 1000 but can vary from unit to unit by as much as ±50% of nominal. It is required to apply negative feedback to this amplifier so that the variability of the closed-loop gain of the resulting feedback amplifier is limited to ±1%. What is the largest possible nominal value of closed-loop gain that can be achieved? Now if three of these feedback amplifiers are placed in cascade, what is the nominal value of the gain of the resulting cascade amplifier? What is the expected variability of this gain?

11.9 The op amp in the circuit of Fig. P11.9 has an open-circuit voltage gain μ, a differential input resistance R_{id}, and a negligibly small output resistance. It is connected in the noninverting configuration with a feedback network consisting of a voltage divider (R_1, R_2). While β is still determined by the divider ratio [i.e., $\beta = R_1/(R_1 + R_2)$], the open-loop gain A is no longer simply equal to μ. This is because the feedback network now loads the input of the amplifier (because of the finite R_{id}). To determine the value of A, use the method outlined in Section 11.1.3 to determine the loop gain $A\beta$. Thus show that

$$A = \mu \frac{R_{id}}{R_{id} + (R_1 \| R_2)}$$

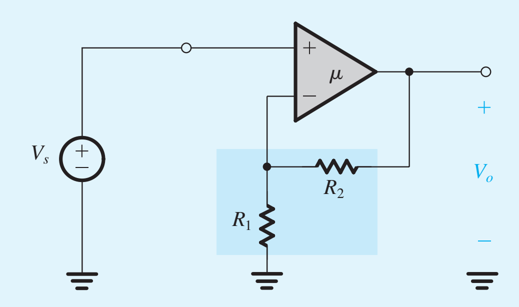


Figure P11.9

Section 11.2: Some Properties of Negative Feedback

11.10 For the negative-feedback loop of Fig. 11.1, find the loop gain $A\beta$ for which the sensitivity of closed-loop gain to open-loop gain [i.e., $(dA_f/A_f)/(dA/A)$] is –40 dB. For what value of $A\beta$ does the sensitivity become 1/5?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

D 11.11 A designer is considering two possible designs of a feedback amplifier. The ultimate goal is $A_f = 10$ V/V. One design employs an amplifier for which $A = 1000$ V/V and the other uses $A = 500$ V/V. Find β and the desensitivity factor in both cases. If the $A = 1000$ amplifier units have a gain uncertainty of $\pm 10\%$, what is the gain uncertainty for the closed-loop amplifiers utilizing this amplifier type? If the same result is to be achieved with the $A = 500$ amplifier, what is the maximum allowable uncertainty in its gain?

D 11.12 A designer is required to achieve a closed-loop gain of $10 \pm 0.1\%$ V/V using a basic amplifier whose gain variation is $\pm 10\%$. What nominal value of A and β (assumed constant) are required?

D 11.13 A circuit designer requires a gain of $25 \pm 1\%$ V/V using an amplifier whose gain varies by a factor of 10 over temperature and time. What is the lowest gain required? The value of β? (*Hint*: Since the change in the open-loop gain is very large, do *not* use differential analysis.)

D 11.14 A power amplifier employs an output stage whose gain varies from 2 to 12 for various reasons. What is the gain of an ideal (nonvarying) amplifier connected to drive it so that an overall gain with feedback of $100 \pm 5\%$ V/V can be achieved? What is the value of β to be used? What are the requirements if A_f must be held within $\pm 0.5\%$? For each of these situations, what preamplifier gain and feedback factor β are required if A_f is to be 10 V/V (with the two possible tolerances)? (*Hint*: Since the change in the open-loop gain is very large, do *not* use differential analysis.)

D 11.15 It is required to design an amplifier with a gain of 100 that is accurate to within $\pm 1\%$. You have available amplifier stages with a gain of 1000 that is accurate to within $\pm 30\%$. Provide a design that uses a number of these gain stages in cascade, with each stage employing negative feedback of an appropriate amount. Obviously, your design should use the lowest possible number of stages while meeting specification.

D *11.16 It is required to design an amplifier to have a nominal closed-loop gain of 10 V/V using a battery-operated amplifier whose gain reduces to half its normal full-battery value over the life of the battery. If only 2% drop in closed-loop gain is desired, what nominal open-loop amplifier gain must be used in the design? (Note that since the change in A is large, it is inaccurate to use differentials.) What value of β should be chosen? If component-value variation in the β network may produce as much as a $\pm 1\%$ variation in β, to what value must A be raised to ensure the required minimum gain?

D 11.17 Design a feedback amplifier that has a closed-loop gain of 100 V/V and is relatively insensitive to change in basic-amplifier gain. In particular, it should provide a reduction in A_f to 99 V/V for a reduction in A to one-tenth its nominal value. What is the required loop gain? What nominal value of A is required? What value of β should be used? What would the closed-loop gain become if A were increased tenfold? If A were made infinite?

11.18 Consider an amplifier having a midband gain A_M and a low-frequency response characterized by a pole at $s = -\omega_L$ and a zero at $s = 0$. Let the amplifier be connected in a negative-feedback loop with a feedback factor β. Find an expression for the midband gain and the lower 3-dB frequency of the closed-loop amplifier. By what factor have both changed?

11.19 A capacitively coupled amplifier has a midband gain of 1000 V/V, a single high-frequency pole at 10 kHz, and a single low-frequency pole at 100 Hz. Negative feedback is employed so that the midband gain is reduced to 10. What are the upper and lower 3-dB frequencies of the closed-loop gain?

D 11.20 Low-cost audio power amplifiers often avoid direct coupling of the loudspeaker to the output stage because any resulting dc bias current in the speaker can use up (and thereby waste) its limited mechanical dynamic range. Unfortunately, the coupling capacitor needed can be large! But feedback helps. For example, for an 8-Ω loudspeaker and $f_L = 100$ Hz, what size capacitor is needed? Now, if feedback is arranged around the amplifier and the speaker so that a closed-loop gain $A_f = 10$ V/V is obtained from an amplifier whose open-loop gain is 1000 V/V, what value of f_{Lf} results? If the ultimate product-design specification requires a 50-Hz cutoff, what capacitor can be used?

D *11.21 It is required to design a dc amplifier with a low-frequency gain of 1000 and a 3-dB frequency of 1 MHz. You have available gain stages with a gain of 1000 but with a dominant high-frequency pole at 20 kHz. Provide a design that employs a number of such stages in cascade, each with negative feedback of an appropriate amount. Use identical stages.

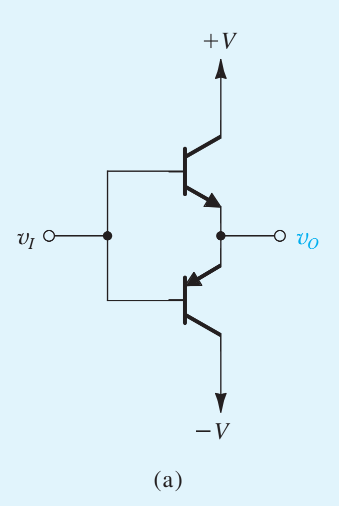
Hint: The 3-dB frequency of a cascade of N identical gain stages, each with a 3-dB frequency $f_{3dB}|_{stage}$ is given by

$$f_{3dB}|_{cascade} = f_{3dB}|_{stage}\sqrt{2^{1/N} - 1}$$

D 11.22 Design a supply-ripple-reduced power amplifier for which the output stage can be modeled by the block diagram of Fig. 11.5, where $A_1 = 0.9$ V/V, and the power-supply ripple $V_N = \pm 1$ V. A closed-loop gain of 10 V/V is desired. What is the gain of the low-ripple preamplifier needed to reduce the output ripple to ±100 mV? To ±10 mV? To ±1 mV? For each case, specify the value required for the feedback factor β.

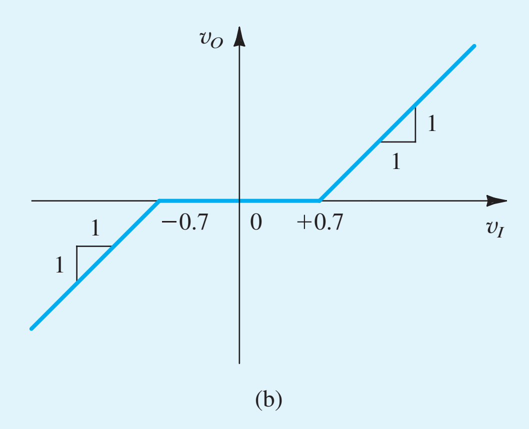
D 11.23 A feedback amplifier is to be designed using a feedback loop connected around a two-stage amplifier. The first stage is a direct-coupled, small-signal amplifier with a high upper 3-dB frequency. The second stage is a power-output stage with a midband gain of 10 V/V and upper and lower 3-dB frequencies of 8 kHz and 80 Hz, respectively. The feedback amplifier should have a midband gain of 100 V/V and an upper 3-dB frequency of 40 kHz. What is the required gain of the small-signal amplifier? What value of β should be used? What does the lower 3-dB frequency of the overall amplifier become?

***11.24** The complementary BJT follower shown in Fig. P11.24(a) has the approximate transfer characteristic shown in Fig. P11.24(b). Observe that for $-0.7\text{ V} \leq v_I \leq +0.7$ V, the output is zero. This "dead band" leads to crossover distortion (see Section 12.3). Consider this follower to be driven by the output of a differential amplifier of gain 100



(a)

Figure P11.24



(b)

Figure P11.24 *continued*

whose positive-input terminal is connected to the input signal source v_S and whose negative-input terminal is connected to the emitters of the follower. Sketch the transfer characteristic v_O versus v_S of the resulting feedback amplifier. What are the limits of the dead band, and what are the gains outside the dead band?

D 11.25 A particular amplifier has a nonlinear transfer characteristic that can be approximated as follows:

(a) For small input signals, $|v_I| \leq 10$ mV, $v_O/v_I = 10^3$.
(b) For intermediate input signals, 10 mV $\leq |v_I| \leq 60$ mV, $\Delta v_O/\Delta v_I = 10^2$.
(c) For large input signals, $|v_I| \geq 60$ mV, the output saturates.

If the amplifier is connected in a negative-feedback loop, find the feedback factor β that reduces the factor-of-10 change in gain (occurring at $|v_I| = 10$ mV) to only a 10% change. What is the transfer characteristic v_O versus v_S of the amplifier with feedback?

Section 11.3: The Feedback Voltage Amplifier

D 11.26 For the feedback voltage amplifier of Fig. 11.8(a), let the op amp have an infinite input resistance, a zero output resistance, and a finite open-loop gain of 1000 V/V. If $R_1 = 10$ kΩ, what value should R_2 have to obtain an ideal closed-loop gain of 10? Now, calculate the loop gain $A\beta$ and use it to find the actual value of the closed-loop gain A_f. If A_f is to be exactly 10, what must the value of R_2 be?

D 11.27 Consider the series–shunt feedback amplifier in Fig. 11.11(a), which is analyzed in Example 11.3.

(a) If $R_1 = 10$ kΩ, find the value of R_2 that results in an ideal closed-loop gain of 10.

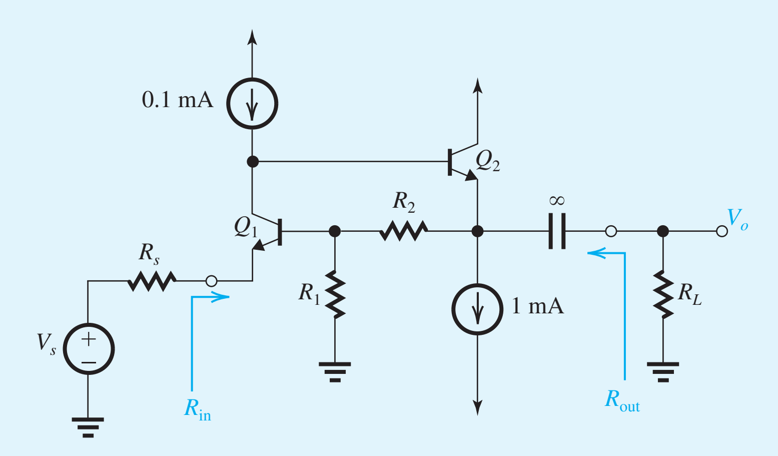


Figure P11.29

(b) Use the expression for $A\beta$ derived in Example 11.3 to find the value of the loop gain for the case $\mu = 1000$, $R_{id} = 100\text{ k}\Omega$, $r_o = 1\text{ k}\Omega$, $R_s = 100\text{ k}\Omega$, and $R_L = 10\text{ k}\Omega$. Hence determine the value of the closed-loop gain A_f.

(c) By what factor must μ be increased to ensure that A_f is within 1% of the ideal value of 10?

D 11.28 Consider the series–shunt feedback amplifier of Fig. 11.8(b) that is analyzed in Example 11.2.

(a) If $R_1 = 1\text{ k}\Omega$, what value should R_2 have to obtain a closed-loop gain whose ideal value is 5 V/V?

(b) If $g_{m1} = g_{m2} = 4\text{ mA/V}$, $R_{D1} = R_{D2} = 10\text{ k}\Omega$, and the MOSFET's r_o is very large, use the expression for $A\beta$ derived in Example 11.2 to find the value of $A\beta$ and hence determine the closed-loop gain A_f.

***11.29** In the series–shunt feedback amplifier shown in Fig. P11.29, the devices operate with $V_{BE} = 0.7$ V and have $\beta_1 = \beta_2 = 100$. The input signal V_s has a zero dc component. Resistances $R_s = 100\ \Omega$, $R_1 = 1\text{ k}\Omega$, $R_2 = 10\text{ k}\Omega$, and $R_L = 1\text{ k}\Omega$.

(a) If the loop gain is large, what do you expect the closed-loop gain to be? Give both an expression and its value.

(b) Find the dc emitter current in each of Q_1 and Q_2. Also, find the dc voltage at the emitter of Q_2.

(c) Calculate the value of the loop gain $A\beta$. (*Hint*: Set $V_s = 0$ and break the loop at the base of Q_1. Simplify the circuit by eliminating dc sources.)

(d) Calculate the value of A_f.

D 11.30 Consider the series–shunt feedback amplifier of Fig. 11.8(c), which was the subject of Exercise 11.6. Assume that the voltage divider (R_1, R_2) is implemented with a 1-MΩ potentiometer. Assume that the MOSFET is biased so that $g_m = 4$ mA/V and r_o is large. Also, $R_D = 10\text{ k}\Omega$. Find the value of R_1 that results in a closed-loop gain of 5 V/V.

D 11.31 Figure P11.31 shows a series–shunt feedback amplifier known as a "feedback triple." All three MOSFETs are biased to operate at $g_m = 4$ mA/V. You may neglect their r_o's.

(a) Select a value for R_F that results in a closed-loop gain that is ideally 10 V/V.

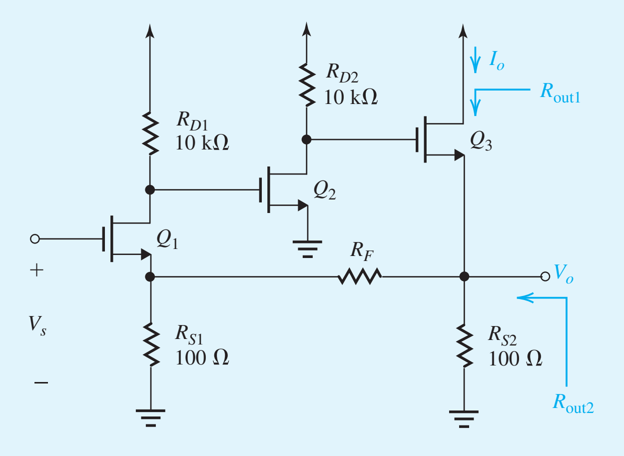


Figure P11.31

(b) Determine the loop gain $A\beta$ and hence the value of A_f. By what percentage does A_f differ from the ideal value you designed for? How can you adjust the circuit to make A_f equal to 10?

11.32 Figure P11.32 shows a series–shunt feedback amplifier without details of the bias circuit.

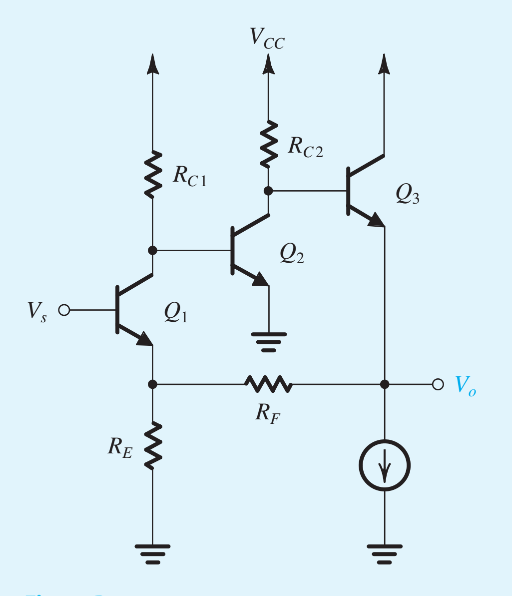


Figure P11.32

(a) If R_E is selected to be 50 Ω, find the value for R_F that results in a closed-loop gain with an ideal value of 25 V/V.

(b) If Q_1 is biased at 1 mA, Q_2 at 2 mA, and Q_3 at 5 mA, and assuming that the transistors have $h_{fe} = 100$ and large r_o, and that $R_{C1} = 2$ kΩ and $R_{C2} = 1$ kΩ, find the value of the loop gain $A\beta$ and hence of the closed-loop gain A_f.

D 11.33 The current-mirror-loaded differential amplifier in Fig. P11.33 has a feedback network consisting of the voltage divider (R_1, R_2), with $R_1 + R_2 = 1$ MΩ. The devices are sized to operate at $|V_{OV}| = 0.2$ V. For all devices, $|V_A| = 10$ V. The input signal source has a zero dc component.

(a) Find the loop gain $A\beta$ and hence the value of A.

(b) Find the values of R_1 and R_2 that result in a closed-loop gain of exactly 5 V/V.

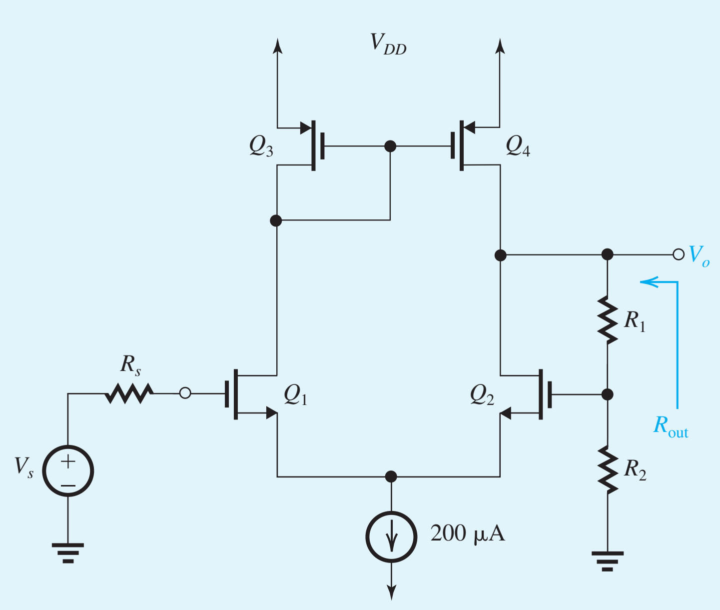


Figure P11.33

Section 11.4: Systematic Analysis of Feedback Voltage Amplifiers (Series–Shunt)

11.34 A series–shunt feedback amplifier employs a basic amplifier with input and output resistances each of 2 kΩ and gain $A = 1000$ V/V. The feedback factor $\beta = 0.1$ V/V. Find

the gain A_f, the input resistance R_{if}, and the output resistance R_{of} of the closed-loop amplifier.

11.35 For a particular amplifier connected in a feedback loop in which the output voltage is sampled, measurement of the output resistance before and after the loop is connected shows a change by a factor of 200. Is the resistance with feedback higher or lower? What is the value of the loop gain $A\beta$? If R_{of} is 100 Ω, what is R_o without feedback?

11.36 The formulas for R_{if} and R_{of} in Eqs. (11.20) and (11.23), respectively, also apply for the case in which A is a function of frequency. In this case, the resulting impedances Z_{if} and Z_{of} will be functions of frequency. Consider the case of a series–shunt amplifier that has an input resistance R_i, an output resistance R_o, and open-loop gain $A = A_0/[1+(s/\omega_H)]$, and a feedback factor β that is independent of frequency. Find Z_{if} and Z_{of} and give an equivalent circuit for each, together with the values of all the elements in the equivalent circuits.

11.37 A feedback amplifier utilizing voltage sampling and employing a basic voltage amplifier with a gain of 1000 V/V and an input resistance of 1000 Ω has a closed-loop input resistance of 10 kΩ. What is the closed-loop gain? If the basic amplifier is used to implement a unity-gain voltage buffer, what input resistance do you expect?

11.38 Consider the noninverting op-amp circuit of Example 11.4 for the case $R_1 = \infty$ and $R_2 = 0$.

(a) What is the value of β, and what is the ideal value of the closed-loop gain?
(b) Adapt the expressions found in Example 11.4 to obtain expressions for A and $A\beta$ for this case.
(c) For $\mu = 10^4$, $R_{id} = 100$ kΩ, $R_s = 10$ kΩ, $r_o = 1$ kΩ, and $R_L = 2$ kΩ, find A, $A\beta$, A_f, R_{in}, and R_{out}.

***11.39** This problem deals with the series–shunt feedback amplifier of Fig. P11.29 and overlaps somewhat with Problem 11.29. Thus, if you have already solved 11.29, you can use some of the results in the solution of this problem. The devices operate with $V_{BE} = 0.7$ V and have $\beta_1 = \beta_2 = 100$. The input signal V_s has a zero dc component. Resistances $R_s = 100$ Ω, $R_1 = 1$ kΩ, $R_2 = 10$ kΩ, and $R_L = 1$ kΩ.

(a) If the loop gain is large, what do you expect the closed-loop gain V_o/V_s to be? Give both an expression and its approximate value.
(b) Find the dc emitter current in each of Q_1 and Q_2. Also find the dc voltage at the emitter of Q_2.
(c) Sketch the A circuit without the dc sources. Derive expressions for A, R_i, and R_o, and find their values.
(d) Give an expression for β and find its value.
(e) Find the closed-loop gain V_o/V_s, the input resistance R_{in}, and the output resistance R_{out}. By what percentage does the value of A_f differ from the approximate value found in (a)?

SIM **D *11.40** Figure P11.40 shows a series–shunt amplifier with a feedback factor $\beta = 1$. The amplifier is designed so that $v_O = 0$ for $v_S = 0$, with small deviations in v_O from 0 V dc being minimized by the negative-feedback action. The technology utilized has $k_n' = 2k_p' = 120\ \mu$A/V^2, $|V_t| = 0.7$ V, and $|V_A'| = 24$ V/μm.

(a) Show that the feedback is negative.
(b) With the feedback loop opened at the gate of Q_2, and the gate terminals of Q_1 and Q_2 grounded, find the dc current and the overdrive voltage at which each of Q_1 to Q_5 is operating. Ignore the Early effect. Also find the dc voltage at the output.
(c) Find g_m and r_o of each of the five transistors.
(d) Find expressions and values of A and R_o. Assume that the bias current sources are ideal.
(e) Find the gain with feedback, A_f, and the output resistance R_{out}.
(f) How would you modify the circuit to realize a closed-loop voltage gain of 5 V/V? What is the value of output resistance obtained?

***11.41** Figure P11.41 shows a series–shunt amplifier in which the three MOSFETs are sized to operate at $|V_{OV}| = 0.2$ V. Let $|V_t| = 0.5$ V and $|V_A| = 10$ V. The current sources utilize single transistors and thus have output resistances equal to r_o.

(a) Show that the feedback is negative.
(b) Assuming the loop gain to be large, what do you expect the closed-loop voltage gain V_o/V_s to be approximately?
(c) If V_s has a zero dc component, find the dc voltages at nodes S_1, G_2, S_3, and G_3. Verify that each of the current sources has the minimum required dc voltage across it for proper operation.
(d) Find the A circuit. Calculate the gain of each of the three stages and the overall voltage gain, A.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

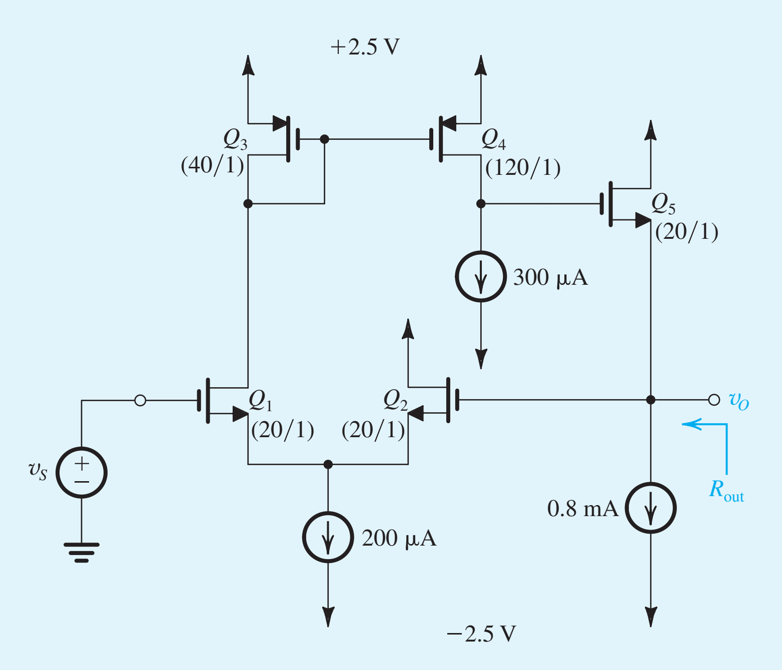


Figure P11.40

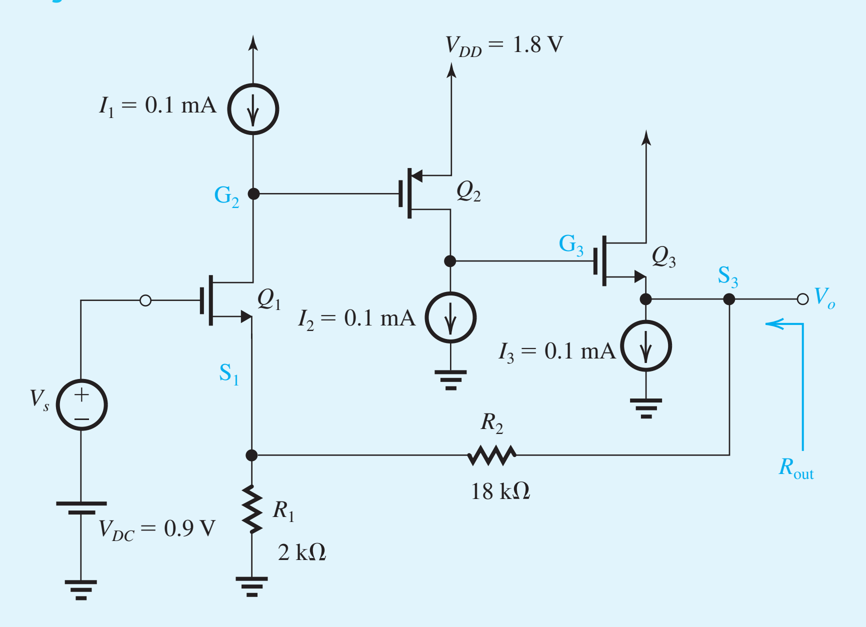


Figure P11.41

[*Hint*: A CS amplifier with a resistance R_s in the source lead has an effective transconductance $g_m/(1+g_mR_s)$ and an output resistance $r_o(1+g_mR_s)$.]

(e) Find β.

(f) Find $A_f = V_o/V_s$. By what percentage does this value differ from the approximate value obtained in (b)?

(g) Find the output resistance R_{out}.

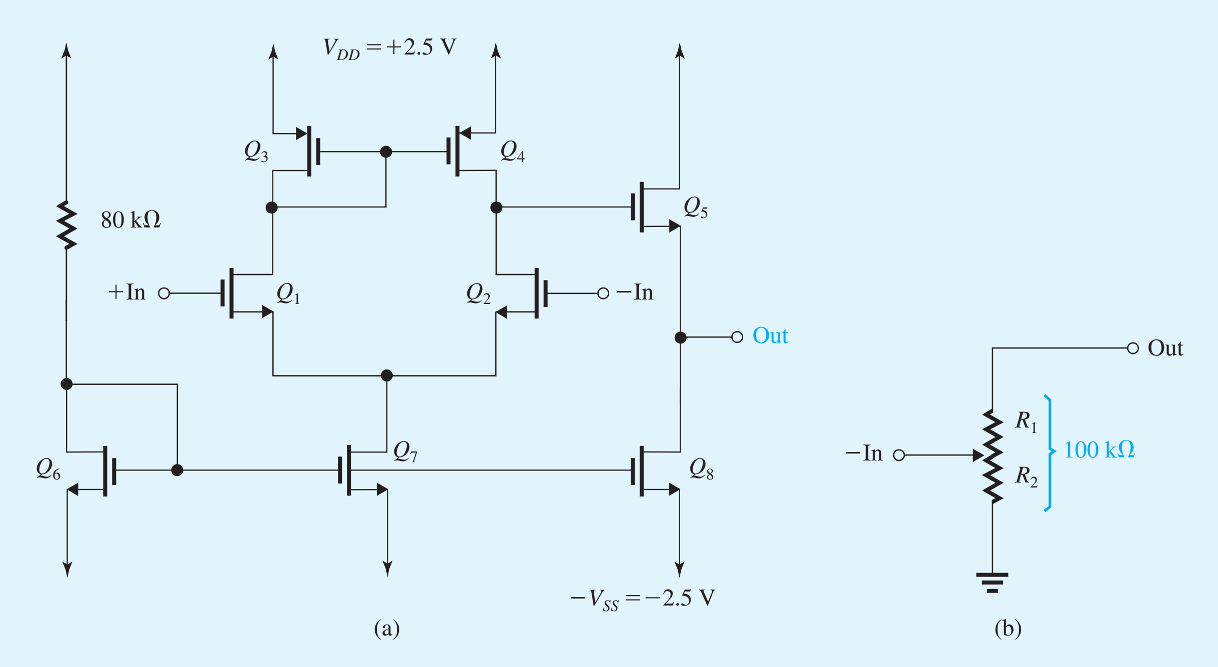


Figure P11.43

D *11.42 This problem deals with the series–shunt feedback amplifier of Fig. P11.33. Certain aspects of this amplifier were considered in Problem 11.33. If you have already solved problem 11.33, you will have the opportunity to compare results. The current-mirror-loaded differential amplifier has a feedback network consisting of the voltage divider (R_1, R_2), with $R_1 + R_2 = 1$ MΩ. The devices are sized to operate at $|V_{OV}| = 0.2$ V. For all devices, $|V_A| = 10$ V. The input signal source has a zero dc component.

(a) Show that the feedback is negative.
(b) What do you expect the dc voltage at the gate of Q_2 to be? At the output? (Neglect the Early effect.)
(c) Find the A circuit. Derive an expression for A and find its value.
(d) Select values for R_1 and R_2 to obtain a closed-loop voltage gain $V_o/V_s = 5$ V/V.
(e) Find the value of R_{out}.
(f) Utilizing the open-circuit, closed-loop gain (5 V/V) and the value of R_{out} found in (e), find the value of gain obtained when a resistance $R_L = 10$ kΩ is connected to the output.
(g) As an alternative approach to (f) above, redo the analysis of the A circuit including R_L. Then utilize the values of R_1 and R_2 found in (d) to determine β and A_f. Compare the value of A_f to that found in (f).

D **11.43 The CMOS op amp in Fig. P11.43(a) is fabricated in a 1-µm technology for which $V_{tn} = -V_{tp} = 0.75$ V, $\mu_n C_{ox} = 2\mu_p C_{ox} = 100\ \mu\text{A/V}^2$, and $V_A' = 10$ V/µm. All transistors in the circuit have $L = 1\ \mu$m.

(a) It is required to perform a dc bias design of the circuit. For this purpose, let the two input terminals be at zero volts dc and neglect channel-length modulation (i.e., let $V_A = \infty$). Design to obtain $I_{D1} = I_{D2} = 50\ \mu$A, $I_{D5} = 250\ \mu$A, and $V_O = 0$, and operate all transistors except for the source follower Q_5 at $V_{OV} = 0.25$ V. Assume that Q_1 and Q_2 are perfectly matched, and similarly for Q_3 and Q_4. For each transistor, find I_D and W/L.
(b) What is the allowable range of input common-mode voltage?
(c) Find g_m for each of Q_1, Q_2, and Q_5.
(d) For each transistor, calculate r_o.
(e) The 100-kΩ potentiometer shown in Fig. 11.43(b) is connected between the output terminal (Out) and the inverting input terminal (–In) to provide negative

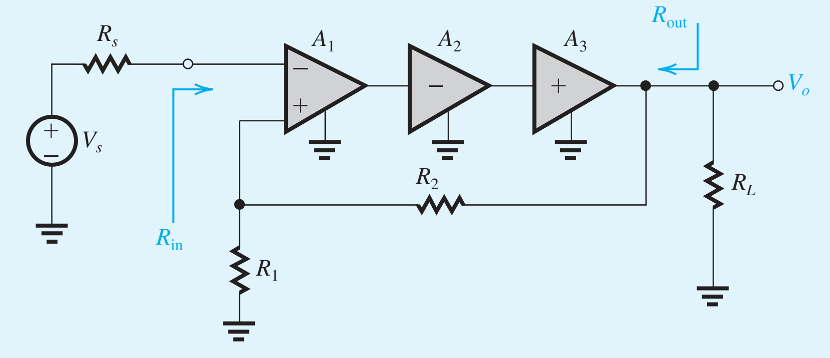


Figure P11.45

feedback whose amount is controlled by the setting of the wiper. A voltage signal V_s is applied between the noninverting input (+In) and ground. A load resistance $R_L = 100$ kΩ is connected between the output terminal and ground. The potentiometer is adjusted to obtain a closed-loop gain $A_f \equiv V_o/V_s \simeq 10$ V/V.

Specify the required setting of the potentiometer by giving the values of R_1 and R_2. Toward this end, find the A circuit (supply a circuit diagram), the value of A, the β circuit (supply a circuit diagram), and the value of β.

(f) What is the output resistance of the feedback amplifier, excluding R_L?

D *11.44 Figure P11.32 shows a series–shunt feedback amplifier without details of the bias circuit.

(a) Eliminating the dc sources, sketch the A circuit and the circuit for determining β.
(b) Show that if $A\beta$ is large then the closed-loop voltage gain is given approximately by

$$A_f \equiv \frac{V_o}{V_s} \simeq \frac{R_F + R_E}{R_E}$$

(c) If R_E is selected equal to 50 Ω, find R_F that will result in a closed-loop gain of approximately 25 V/V.
(d) If Q_1 is biased at 1 mA, Q_2 at 2 mA, and Q_3 at 5 mA, and assuming that the transistors have $h_{fe} = 100$, find approximate values for R_{C1} and R_{C2} to obtain gains from the stages of the A circuit as follows: a voltage gain of Q_1 of about –10 and a voltage gain of Q_2 of about –50.
(e) For your design, what is the closed-loop voltage gain realized?
(f) Calculate the input and output resistances of the closed-loop amplifier designed.

D *11.45 Figure P11.45 shows a three-stage feedback amplifier:

A_1 has an 82-kΩ differential input resistance, a 20-V/V open-circuit differential voltage gain, and a 3.2-kΩ output resistance.

A_2 has a 5-kΩ input resistance, a 20-mA/V short-circuit transconductance, and a 20-kΩ output resistance.

A_3 has a 20-kΩ input resistance, unity open-circuit voltage gain, and a 1-kΩ output resistance.

The feedback amplifier feeds a 1-kΩ load resistance and is fed by a signal source with a 9-kΩ resistance.

(a) Show that the feedback is negative.
(b) If $R_1 = 20$ kΩ, find the value of R_2 that results in a closed-loop gain V_o/V_s that is ideally 5 V/V.
(c) Supply the small-signal equivalent circuit.
(d) Sketch the A circuit and determine A.
(e) Find β and the amount of feedback.
(f) Find the closed-loop gain $A_f \equiv V_o/V_s$.
(g) Find the feedback amplifier's input resistance R_{in}.
(h) Find the feedback amplifier's output resistance R_{out}.
(i) If the high-frequency response of the open-loop gain A is dominated by a pole at 100 Hz, what is the upper 3-dB frequency of the closed-loop gain?
(j) If for some reason A_1 drops to half its nominal value, what is the percentage change in A_f?

Section 11.5: Other Feedback-Amplifier Types

D 11.46 Refer to the circuit in Fig. 11.17(a), which is analyzed in Example 11.6. Select a value for R_F that results in a closed-loop transconductance $A_f \equiv I_o/V_s \simeq 10$ mA/V. Use the formulas derived in Example 11.6 to find the actual value of A_f realized. Let $\mu = 1000$, $R_{id} = 100$ kΩ, $g_m = 2$ mA/V, and $r_{o2} = 20$ kΩ.

D 11.47 Figure P11.47 shows a feedback current amplifier. The feedback network consists of the highlighted two-port network comprising R_M and R_F. It is fed with the output current I_o and delivers a feedback current I_f at its port 1 to the input node. The feedback factor β is the current ratio I_f/I_o measured with port 1 short-circuited (because it is connected in shunt with the amplifier input).

(a) Find an expression for β and hence for the ideal value of $A_f \equiv I_o/I_s$.

(b) Setting $I_s = 0$, break the loop at the gate of Q_2 and thus determine the loop gain $A\beta$. Show that

$$A = -\frac{g_{m2}R_D}{1 + 1/[g_{m1}(R_M + R_F)]}$$

(c) For $g_{m1} = g_{m2} = 4$ mA/V, $R_D = 10$ kΩ, and $(R_M + R_F) = 1$ kΩ, find the value of R_M that results in a closed-loop current gain of 5 A/A.

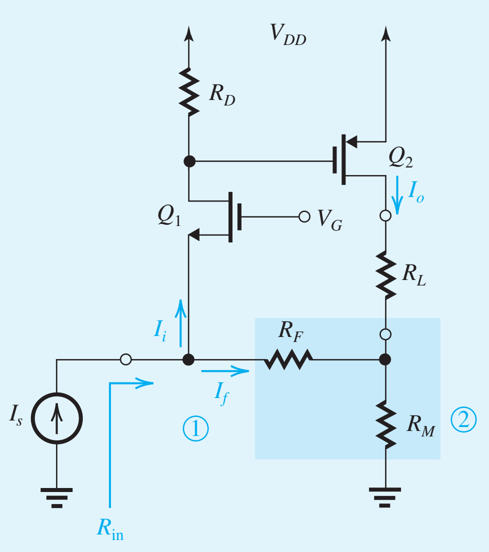
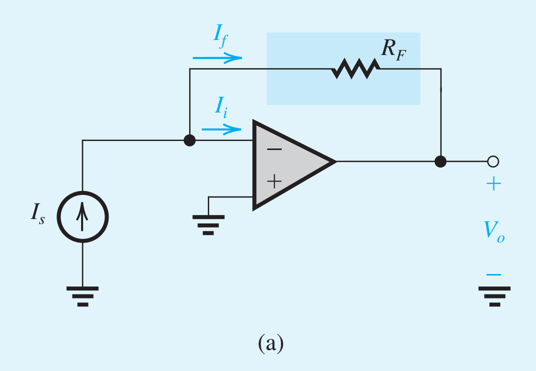


Figure P11.47

D 11.48 Figure P11.48(a) shows a feedback transresistance amplifier formed by an op amp and a feedback resistance R_F. Assume that the op amp is modeled by an input resistance R_{id}, an open-circuit voltage gain μ, and an output resistance r_o.



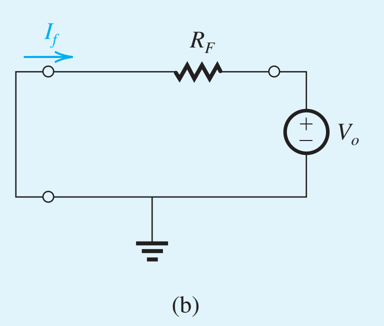


Figure P11.48

(a) Show that the feedback factor β, determined as shown in Fig. P11.48(b), is given by $\beta = -1/R_F$. Hence find the ideal value of the closed-loop gain $A_f \equiv V_o/I_s$. Find R_F that results in A_f of approximately 1 kΩ.

(b) By setting $I_s = 0$ and breaking the loop at the input terminals of the op amp, show that the loop gain is given by

$$A\beta = \mu \frac{R_{id}}{R_{id} + R_F + r_o}$$

(c) For $\mu = 1000$, $R_{id} = 100$ kΩ, $r_o = 1$ kΩ, and R_F having the value found in (a), what is the actual value of A_f realized?

Feedback Transconductance Amplifiers (Series–Series)

11.49 A series–series feedback amplifier employs a transconductance amplifier having a short-circuit transconductance G_m of 0.6 A/V, input resistance of 10 kΩ, and output resistance of 100 kΩ. The feedback network has $\beta = 200$ Ω,

an input resistance (with port 1 open-circuited) of 200 Ω, and an input resistance (with port 2 open-circuited) of 10 kΩ. The amplifier operates with a signal source having a resistance of 10 kΩ and with a load resistance of 10 kΩ. Find A_f, R_{in}, and R_{out}.

11.50 Reconsider the circuit in Fig. 11.21(a), analyzed in Example 11.8, this time with the output voltage taken at the emitter of Q_3. In this case the feedback can be considered to be of the series–shunt type. Note that R_{E2} should now be considered part of the basic amplifier and not of the feedback network.

(a) Determine β.
(b) Find an approximate value for $A_f \equiv V_{e3}/V_s$ assuming that the loop gain remains large (a safe assumption, since the loop in fact does not change).
[*Note*: If you continue with the feedback analysis, you'll find that $A\beta$ in fact changes somewhat; this is a result of the different approximations made in the feedback analysis approach.]
(c) If the loop gain remains at the value calculated in Example 11.8 (i.e., 246.3), find the output resistance R_{out} (measured between the emitter of Q_3 and ground). (Neglect the effect of r_{03}.)

D *11.51 Figure P11.31 (page 851) shows a feedback triple utilizing MOSFETs. All three MOSFETs are biased and sized to operate at $g_m = 4$ mA/V. You may neglect their r_o's (except for the calculation of R_{out1} as indicated below).

(a) Considering the feedback amplifier as a transconductance amplifier with output current I_o, find the value of R_F that results in a closed-loop transconductance of approximately 100 mA/V.
(b) Sketch the A circuit and find the value of $A \equiv I_o/V_i$.
(c) Find $1 + A\beta$ and $A_f \equiv I_o/V_s$. Compare to the value of A_f you designed for. What is the percentage difference? What resistance can you change to make A_f exactly 100 mA/V, and in which direction (increase or decrease)?
(d) Assuming that $r_{o3} = 20$ kΩ, find R_o of the A circuit. For this purpose, recall that the resistance looking into the drain of a MOSFET having a resistance R_s in its source is $(r_o + R_s + g_m r_o R_s)$. Hence find the output resistance R_{out1}. Since the current sampled by the feedback network is exactly equal to the output current, you can use the feedback formula.
(e) If the voltage V_o is taken as the output, in which case the amplifier becomes series–shunt feedback, what is the value of the closed-loop voltage gain V_o/V_s? Assume that R_F has the original value you selected in (a). Note that in this case R_{S2} should be considered part of the amplifier and not the feedback network. The feedback analysis will reveal that $A\beta$ changes somewhat, which may be puzzling given that the feedback loop did not change. The change is due to the different approximation used.
(f) What is the closed-loop output resistance R_{out2} of the voltage amplifier in (e) above?

11.52 Consider the circuit in Fig. P11.52 as a transconductance amplifier with input V_s and output I_o. The transistor is specified in terms of its g_m and r_o.

(a) Sketch the small-signal equivalent circuit using the hybrid-π model of the MOSFET and convince yourself that the feedback circuit is comprised of resistor R_F.
(b) Find the A circuit and the β circuit.
(c) Derive expressions for A, β, $(1 + A\beta)$, A_f, R_o, and R_{of}.

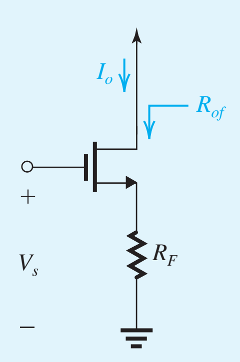


Figure P11.52

D 11.53 The transconductance amplifier in Fig. P11.53 utilizes a differential amplifier with gain μ and a very high input resistance. The differential amplifier drives a transistor Q characterized by its g_m and r_o. A resistor R_F senses the output current I_o.

(a) For $A\beta \gg 1$, find an approximate expression for the closed-loop transconductance $A_f \equiv I_o/V_s$. Hence, select a value for R_F that results in $A_f \simeq 5$ mA/V.
(b) Find the A circuit and derive an expression for A. Evaluate A for the case $\mu = 1000$ V/V, $g_m = 2$ mA/V, $r_o = 20$ kΩ, and the value of R_F you selected in (a).
(c) Give an expression for $A\beta$ and evaluate its value and that of $1 + A\beta$.

(d) Find the closed-loop gain A_f and compare to the value you designed for in (a) above.

(e) Find expressions and values for R_o and R_{of}. [*Hint*: The resistance looking into the drain of a MOSFET with a resistance R_s in its source is $(r_o + R_s + g_m r_o R_s)$.]

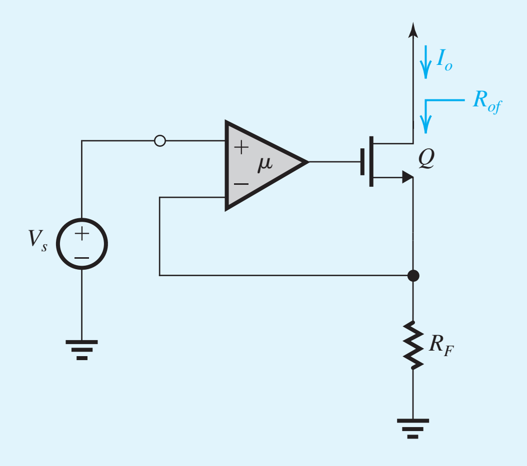


Figure P11.53

***11.54** It is required to show that the output resistance of the BJT circuit in Fig. P11.54 is given by

$$R_o = r_o + [R_e \| (r_\pi + R_b)]\left(1 + g_m r_o \frac{r_\pi}{r_\pi + R_b}\right)$$

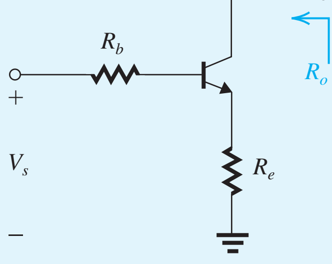
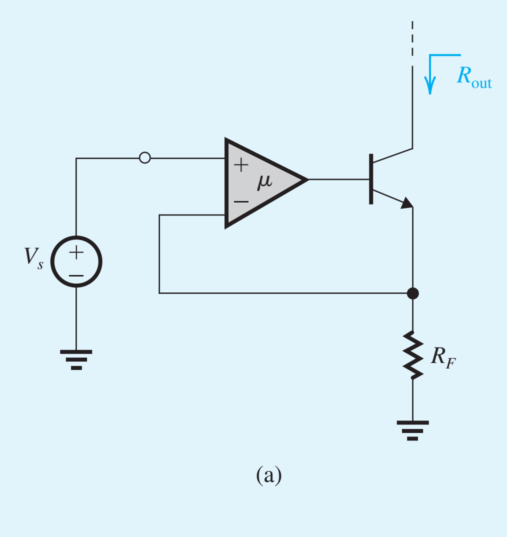


Figure P11.54

To derive this expression, set $V_s = 0$, replace the BJT with its small-signal, hybrid-π model, apply a test voltage V_x to the collector, and find the current I_x drawn from V_x and hence R_o as V_x/I_x. Note that the bias arrangement is not shown. For the case of $R_b = 0$, find the maximum possible value for R_o. Note that this theoretical maximum is obtained when R_e is so large that the signal current in the emitter is nearly zero. In this case, with V_x applied and $V_s = 0$, what is the current in the base, in the $g_m V_\pi$ generator, and in r_o, all in terms of I_x? Show these currents on a sketch of the equivalent circuit with R_e set to ∞.

11.55 As we found out in Example 11.8, whenever the feedback network senses the emitter current of the BJT, the feedback output resistance formula cannot predict the output resistance looking into the collector. To understand this issue more clearly, consider the feedback transconductance amplifier shown in Fig. P11.55(a). To determine the output resistance, we set $V_s = 0$ and apply a test voltage V_x to the collector, as shown in Fig. P11.55(b). Now, let μ be increased to the point where the feedback signal across R_F almost equals the input to the positive terminal of the differential amplifier,



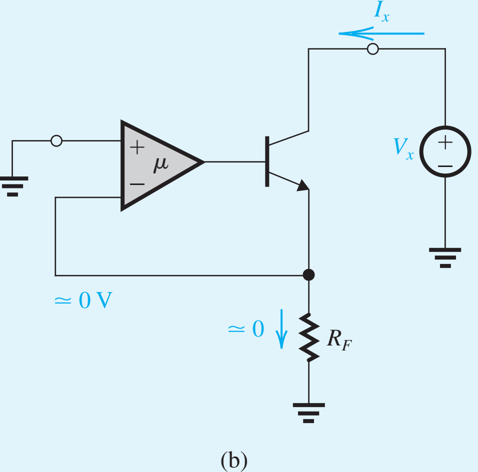


Figure P11.55

now zero. Thus the signal current through R_F will be almost zero. By replacing the BJT with its hybrid-π model, show that

$$R_{out} = r_\pi + (h_{fe} + 1)r_o \simeq h_{fe} r_o$$

where h_{fe} is the transistor β. Thus for large amounts of feedback, R_{out} is limited to a maximum of $h_{fe}r_o$ independent of the amount of feedback. This phenomenon does *not* occur in the MOSFET version of this circuit, where the output resistance can be theoretically made infinite.

11.56 For the feedback transconductance amplifier of Fig. P11.56 derive expressions for A, β, $A\beta$, A_f, R_o, and R_{of}. Evaluate A_f and R_{of} for the case of $g_{m1} = g_{m2} = 4$ mA/V, $R_D = 20$ kΩ, $r_{o2} = 20$ kΩ, $R_F = 100$ Ω, and $R_L = 1$ kΩ. For simplicity, neglect r_{o1} and take r_{o2} into account only when calculating output resistances.

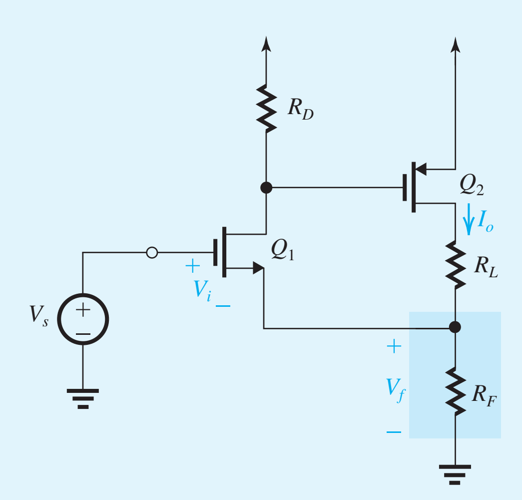


Figure P11.56

D 11.57 For the feedback transconductance amplifier in Fig. P11.57, derive an approximate expression for the closed-loop transconductance $A_f \equiv I_o/V_s$ for the case of $A\beta \gg 1$. Hence select a value for R_2 to obtain $A_f = 100$ mA/V. If Q is biased to obtain $g_m = 1$ mA/V, specify the value of the gain μ of the differential amplifier to obtain an amount of feedback of 60 dB. If Q has $r_o = 50$ kΩ, find the output resistance R_{out}. [*Hint*: Recall that for a MOSFET with a resistance R_s in its source, the resistance looking into the drain is $(r_o + R_s + g_m r_o R_s)$.]

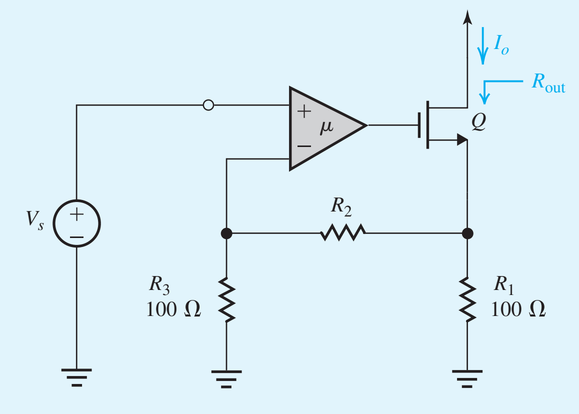


Figure P11.57

SIM **11.58** All the MOS transistors in the feedback transconductance amplifier (series–series) of Fig. P11.58 are sized to operate at $|V_{OV}| = 0.2$ V. For all transistors, $|V_t| = 0.4$ V and $|V_A| = 20$ V.

(a) If V_s has a zero dc component, find the dc voltage at the output, at the drain of Q_1, and at the drain of Q_2.
(b) Find an approximate expression and value for $A_f \equiv I_o/V_s$ for the case $A\beta \gg 1$.
(c) Use feedback analysis to obtain a more precise value for A_f.
(d) Find the value of R_{out}.
(e) If the voltage at the source of Q_5 is taken as the output, find the voltage gain using the value of I_o/V_s obtained in (c). Also find the output resistance of this series–shunt voltage amplifier.

11.59 By setting $V_s = 0$ and breaking the feedback loop, show that the loop gain of the amplifier circuit in Fig. P11.58 is

$$A\beta = g_{m1,2}(r_{o2} \| r_{o4}) \frac{R_F \| r_{o5}}{(R_F \| r_{o5}) + 1/g_{m5}}$$

where $g_{m1,2}$ is the g_m of each of Q_1 and Q_2.

Feedback Transresistance Amplifiers (Shunt–Shunt)

11.60 For the transresistance amplifier analyzed in Example 11.9, use the formulas derived there to evaluate A_f, R_{in}, and R_{out} when μ is one-tenth the value used in the example. That is, evaluate for $\mu = 10^3$ V/V, $R_{id} = \infty$, $r_o = 100$ Ω, $R_F = 10$ kΩ, and $R_s = R_L = 1$ kΩ. Compare to the corresponding values obtained in Example 11.9.

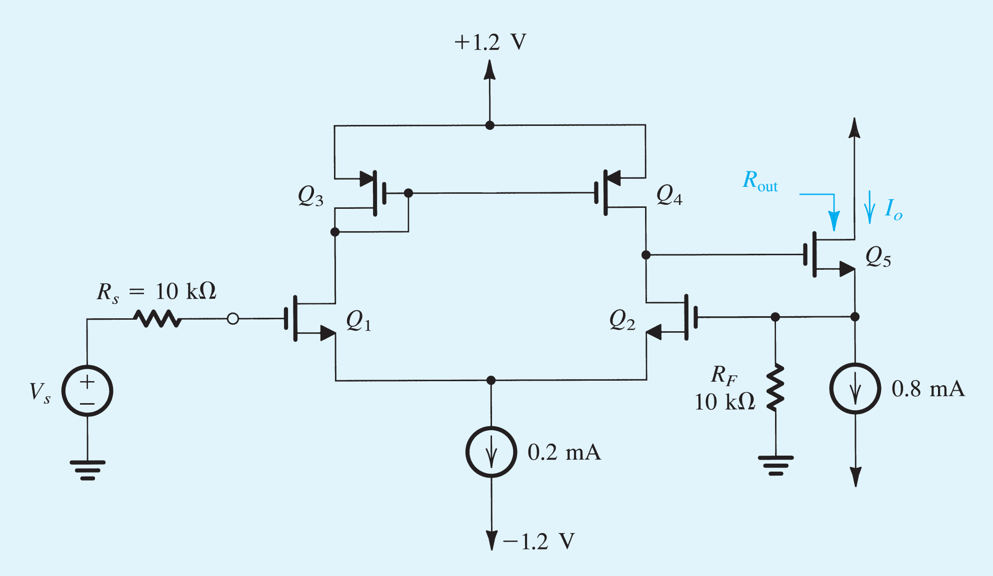


Figure P11.58

11.61 Use the formulas derived in Example 11.9 to solve the problem in Exercise 11.19. Show that the results are identical to those given in the answer to Exercise 11.19.

11.62 By setting $I_s = 0$, replacing the MOSFET with its hybrid-π model, and breaking the feedback loop, determine the loop gain of the feedback amplifier in Fig. E11.19. Hence find the open-loop gain. Evaluate $A\beta$, β, A, and A_f for the numerical values given in Exercise 11.8. Why do the results differ somewhat from those given in the answer to Exercise 11.19?

11.63 The CE BJT amplifier in Fig. P11.63 employs shunt–shunt feedback: Feedback resistor R_F senses the output voltage V_o and provides a feedback current to the base node.

(a) If V_s has a zero dc component, find the dc collector current of the BJT. Assume the transistor $\beta = 100$.
(b) Find the small-signal equivalent circuit of the amplifier with the signal source represented by its Norton equivalent (as we usually do when the feedback connection at the input is shunt).
(c) Find the A circuit and determine the value of A, R_i, and R_o.
(d) Find β and hence $A\beta$ and $1 + A\beta$.
(e) Find A_f, R_{if}, and R_{of} and hence R_{in} and R_{out}.
(f) What voltage gain V_o/V_s is realized? How does this value compare to the ideal value obtained if the loop gain is very large and thus the signal voltage at the base becomes almost zero (like what happens in an inverting op-amp circuit). Note that this single-transistor poor-man's op amp is not that bad!

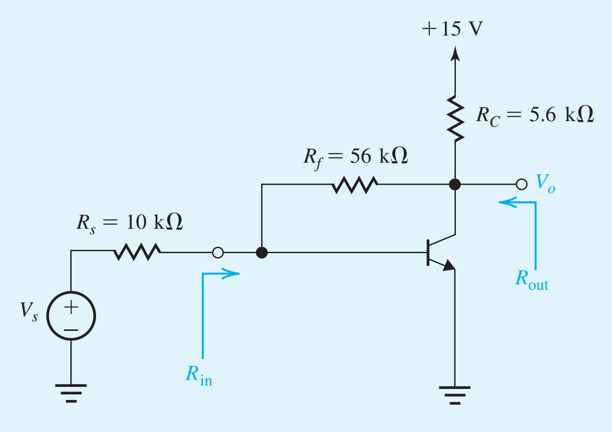


Figure P11.63

D **11.64** The circuit in Fig. P11.64 utilizes a voltage amplifier with gain μ in a shunt–shunt feedback topology with the feedback network composed of resistor R_F. In order to be able to use the feedback equations, you should first convert the signal source to its Norton representation. You will then see that all the formulas derived in Example 11.9 apply here as well.

(a) If the loop gain is very large, what approximate closed-loop voltage gain V_o/V_s is realized? If $R_s = 2\text{ k}\Omega$, give the value of R_F that will result in $V_o/V_s \simeq -10$ V/V.
(b) If the amplifier μ has a dc gain of 10^3 V/V, an input resistance $R_{id} = 100\text{ k}\Omega$, and an output resistance $r_o = 2\text{ k}\Omega$, find the actual V_o/V_s realized. Also find R_{in} and R_{out} (indicated on the circuit diagram). You may use formulas derived in Example 11.9.
(c) If the amplifier μ has an upper 3-dB frequency of 1 kHz and a uniform −20-dB/decade gain rolloff, what is the 3-dB frequency of the gain $|V_o/V_s|$?

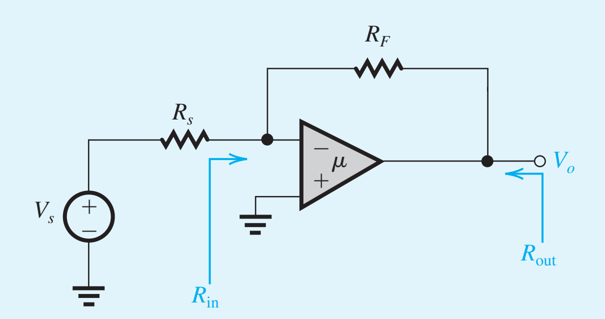


Figure P11.64

11.65 The feedback transresistance amplifier in Fig. P11.65 utilizes two identical MOSFETs biased by ideal current sources $I = 0.4$ mA. The MOSFETs are sized to operate at $V_{OV} = 0.2$ V and have $V_t = 0.5$ V and $V_A = 16$ V. The feedback resistance $R_F = 10\text{ k}\Omega$.

(a) If I_s has a zero dc component, find the dc voltage at the input, at the drain of Q_1, and at the output.
(b) Find g_m and r_o of Q_1 and Q_2.
(c) Provide the A circuit and derive an expression for A in terms of g_{m1}, r_{o1}, g_{m2}, r_{o2}, and R_F.
(d) What is β? Give an expression for the loop gain $A\beta$ and the amount of feedback $(1 + A\beta)$.
(e) Derive an expression for A_f.
(f) Derive expressions for R_i, R_{in}, R_o, and R_{out}.
(g) Evaluate A, β, $A\beta$, A_f, R_i, R_o, R_{in}, and R_{out} for the component values given.

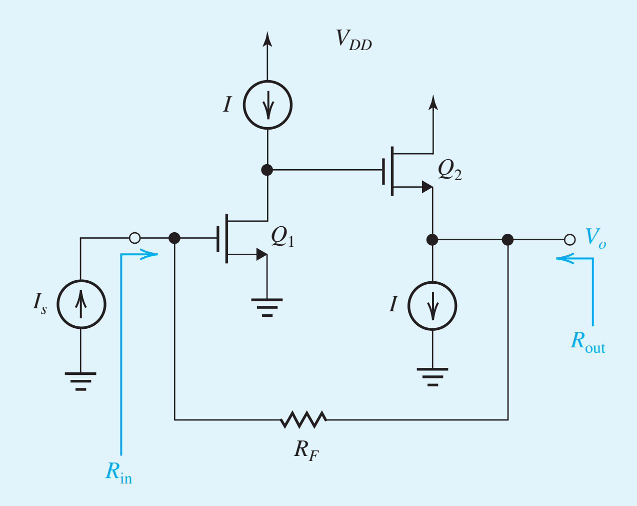


Figure P11.65

11.66 By setting $I_s = 0$ and breaking the feedback loop, find the loop gain of the feedback amplifier in Fig. P11.65. If you have already solved Problem 11.65, compare results. Which result do you think is more accurate, and why? For the numerical values given in Problem 11.65, by how much (in percent) do the two values of loop gain differ?

11.67 Analyze the circuit in Fig. E11.19 from first principles (i.e., do not use the feedback approach) and hence show that

$$A_f \equiv \frac{V_o}{I_s} = -\frac{(R_s \| R_F)\left(g_m - \dfrac{1}{R_F}\right)(r_o \| R_F)}{1 + (R_s \| R_F)\left(g_m - \dfrac{1}{R_F}\right)(r_o \| R_F)/R_F}$$

Comparing this expression to the one given in Exercise 11.19, part (b), you will note that the only difference is that g_m has been replaced by $(g_m - 1/R_F)$. Note that $-1/R_F$ represents the forward transmission in the feedback network, which the feedback-analysis method neglects. What is the condition then for the feedback-analysis method to be reasonably accurate for this circuit?

D 11.68 For the feedback amplifier in Fig. P11.68, select a value for R_F that results in a closed-loop gain $A_f \equiv V_o/I_s \simeq -10$ kΩ. Then, analyze the circuit to determine the actual value of A_f realized. As well, determine R_{in} and R_{out}. Transistors Q_1 and Q_2 are operated so that $g_{m1} = g_{m2} = 4$ mA/V and r_{o1} and r_{o2} can be neglected. Also, $R_{D1} = R_{D2} = 10$ kΩ.

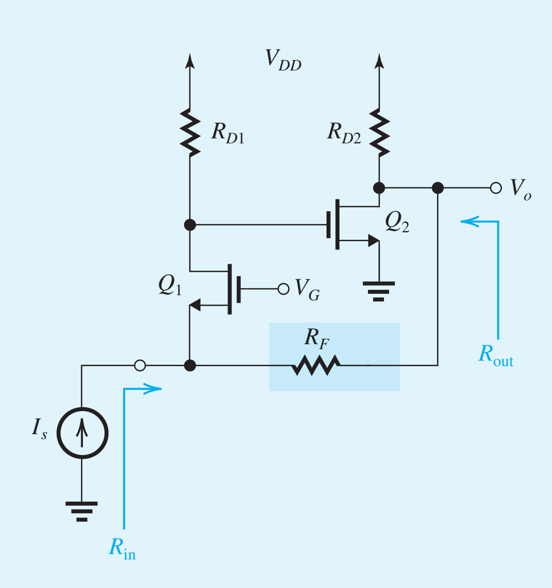


Figure P11.68

11.69 For the feedback transresistance amplifier in Fig. P11.69, let $V_{CC} = -V_{EE} = 5$ V, $R_C = R_E = R_F = 10$ kΩ. The transistors have $V_{BE} = 0.7$ V and $\beta = 100$.

(a) If I_s has a zero dc component, show that Q_1 and Q_2 are operating at dc collector currents of approximately 0.35 mA and 0.58 mA, respectively. What is the dc voltage at the output?

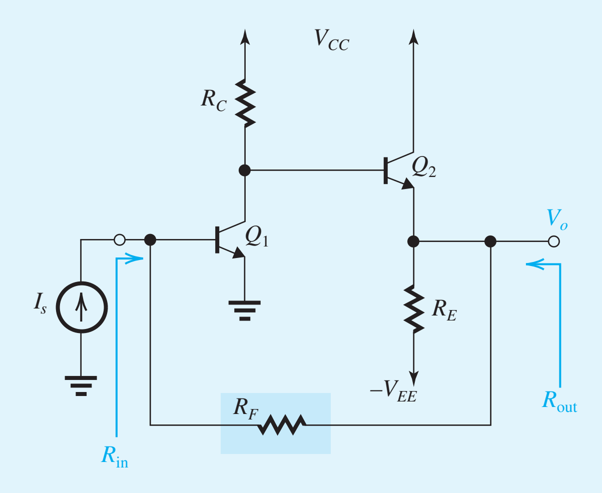


Figure P11.69

(b) Find the A circuit and the value of A, R_i, and R_o. Neglect r_{o1} and r_{o2}.

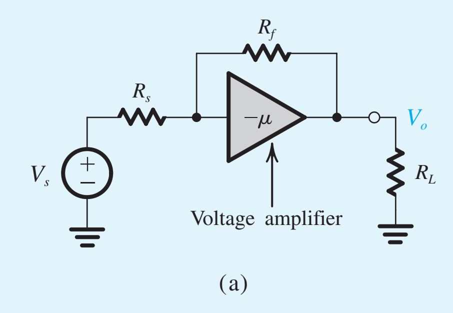
(c) Find the value of β, the loop gain, and the amount of feedback.

(d) Find $A_f \equiv V_o/I_s$, the input resistance R_{in}, and the output resistance R_{out}.

D **11.70 (a) Show that for the circuit in Fig. P11.70(a), if the loop gain is large, the voltage gain V_o/V_s is given approximately by

$$\frac{V_o}{V_s} \simeq -\frac{R_f}{R_s}$$

(b) Using three cascaded stages of the type shown in Fig. P11.70(b) to implement the amplifier μ, design a feedback amplifier with a voltage gain of approximately –100 V/V. The amplifier is to operate between a source resistance $R_s = 10$ kΩ and a load resistance $R_L = 1$ kΩ. Calculate the actual value of V_o/V_s realized, the input resistance (excluding R_s), and the output resistance (excluding R_L). Assume that the BJTs have h_{fe} of 100. [*Note:* In practice, the three amplifier stages are not made identical, for stability reasons.]



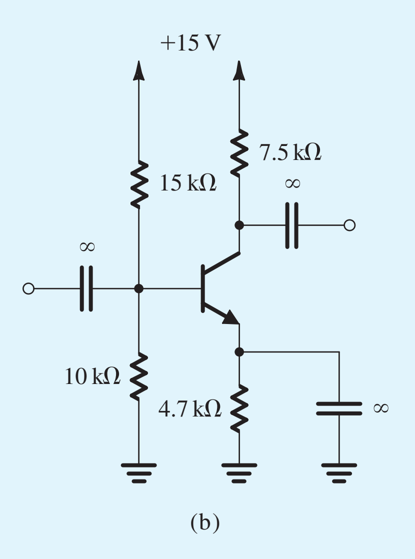


Figure P11.70

D 11.71 Negative feedback is to be used to modify the characteristics of a particular amplifier for various purposes. Identify the feedback topology to be used if:

(a) input resistance is to be lowered and output resistance raised.
(b) both input and output resistances are to be raised.
(c) both input and output resistances are to be lowered.

11.72 The feedback amplifier of Fig. P11.72 consists of a common-gate amplifier formed by Q_1 and R_D, and a feedback circuit formed by the capacitive divider (C_1, C_2) and the common-source transistor Q_f. Note that the bias circuit for Q_f is not shown. It is required to derive expressions for $A_f \equiv V_o/I_s$, R_{in}, and R_{out}. Assume that C_1 and C_2 are sufficiently small that their loading effect on the basic amplifier can be neglected. Also neglect r_o. Find the values of A_f, R_{in}, and R_{out} for the case in which $g_{m1} = 5$ mA/V, $R_D = 10$ kΩ, $C_1 = 0.9$ pF, $C_2 = 0.1$ pF, and $g_{mf} = 2$ mA/V.

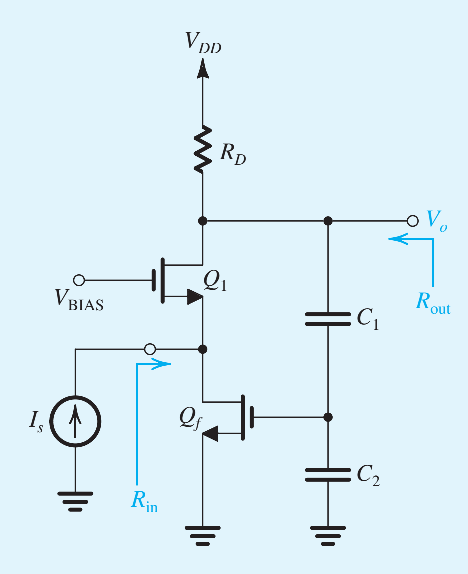


Figure P11.72

D *11.73 Figure P11.73 shows a shunt–shunt feedback amplifier. The MOSFETs have $V_{tn} = 0.6$ V, $V_A = 20$ V, and $\mu_n C_{ox} = 200\ \mu\text{A/V}^2$. The power supply $V_{DD} = 3.3$ V, and $R_L = 2$ kΩ. The coupling capacitor C_C can be assumed to be very large.

(a) Perform a dc design to meet the following specifications: $I_{D1} = 100\ \mu$A, $I_{D2} = 1$ mA, $I_{R2,R1} = 10\ \mu$A, $V_{OV1} = V_{OV2} = 0.2$ V. Neglect the Early effect. Specify the values required for I_1, R_1, R_2, $(W/L)_1$, and $(W/L)_2$.
(b) Find an expression for β and hence an expression for the ideal value of V_o/V_s.
(c) Find the value of R_s that results in V_o/V_s being ideally -6 V/V.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

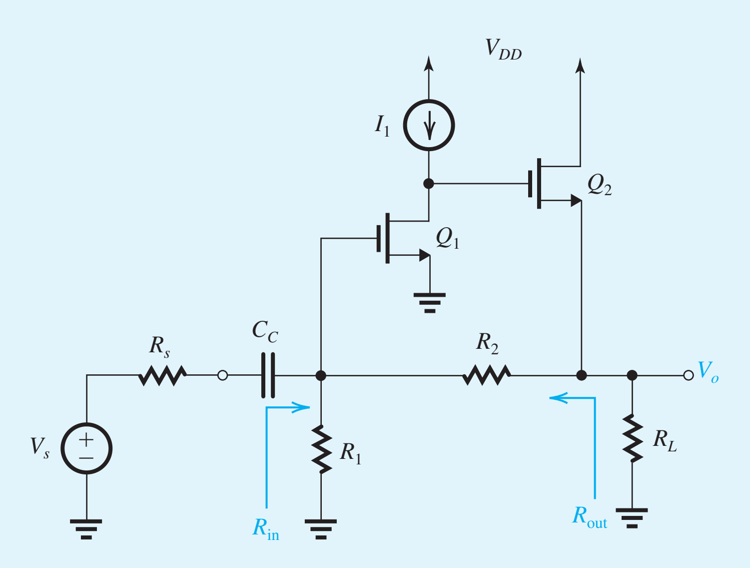


Figure P11.73

(d) Find the A circuit and use it to determine the values of A, R_i, and R_o.
(e) Find the value obtained for V_o/V_s.
(f) Find R_{in} and R_{out}.

Feedback Current Amplifiers (Shunt–Series)

11.74 For the feedback current amplifier in Fig. P11.47:

(a) Provide the A circuit and derive expressions for R_i and A. Neglect r_o of both transistors.
(b) Provide the β circuit and an expression for β.
(c) Find an expression for $A\beta$.
(d) For $g_{m1} = g_{m2} = 5$ mA/V, $R_D = 20$ kΩ, $R_M = 10$ kΩ, and $R_F = 90$ kΩ, find the values of A, β, $A\beta$, A_f, R_i, and R_{if}.
(e) If $r_{o2} = 20$ kΩ and $R_L = 1$ kΩ, find the output resistance as seen by R_L.

D 11.75 Design the feedback current amplifier of Fig. 11.27(a) to meet the following specifications:

(i) $A_f \equiv I_o/I_s = -100$ A/A
(ii) amount of feedback $\simeq 40$ dB
(iii) $R_{in} \simeq 1$ kΩ

Specify the values of R_1, R_2, and μ. Assume that the amplifier μ has infinite input resistance and that $R_s = \infty$. For the MOSFET, $g_m = 5$ mA/V and $r_o = 20$ kΩ. What R_{out} is obtained?

11.76 Consider the feedback current amplifier in Fig. 11.27(a) (which was analyzed in Example 11.10). Let $R_s = R_{id} = \infty$. By setting $I_s = 0$ and breaking the feedback loop at the gate of Q, find an expression for the loop gain $A\beta$. Evaluate $A\beta$ for the component values given in Example 11.10 and hence determine A and A_f. Why do the results differ somewhat from those found in Example 11.10?

11.77 The feedback current amplifier in Fig. P11.77 utilizes two identical NMOS transistors sized so that at $I_D = 0.2$ mA they operate at $V_{OV} = 0.2$ V. Both devices have $V_t = 0.5$ V and $V_A = 10$ V.

(a) If I_s has zero dc component, show that both Q_1 and Q_2 are operating at $I_D = 0.2$ mA. What is the dc voltage at the input?
(b) Find g_m and r_o for each of Q_1 and Q_2.
(c) Find the A circuit and the value of R_i, A, and R_o.
(d) Find the value of β.
(e) Find $A\beta$ and A_f.
(f) Find R_{in} and R_{out}.

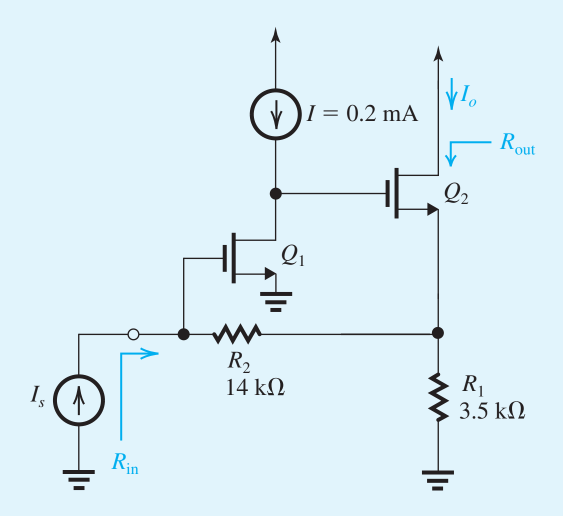


Figure P11.77

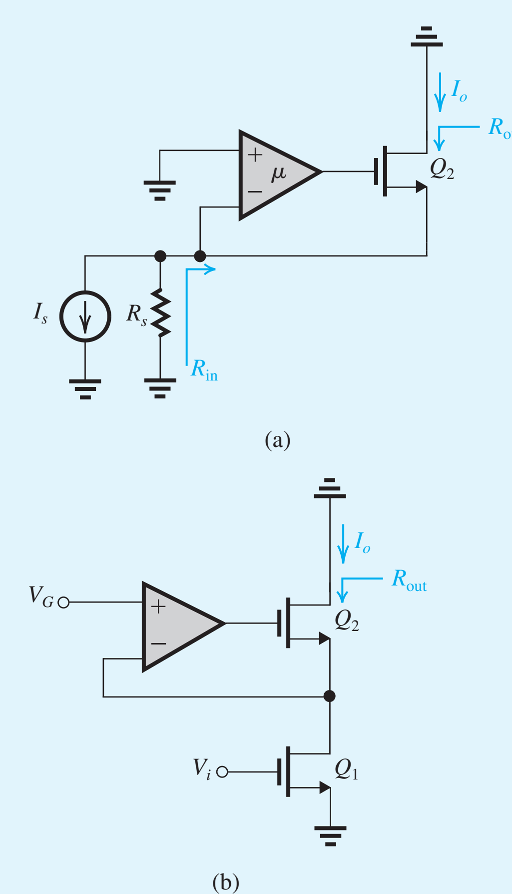


Figure P11.78

***11.78** The feedback current amplifier in Fig. P11.78(a) can be thought of as a "super" CG transistor. Note that rather than connecting the gate of Q_2 to signal ground, an amplifier is placed between source and gate.

(a) If μ is very large, what is the signal voltage at the input terminal? What is the input resistance? What is the current gain I_o/I_s?
(b) For finite μ but assuming that the input resistance of the amplifier μ is very large, find the A circuit and derive expressions for A, R_i, and R_o.
(c) What is the value of β?
(d) Find $A\beta$ and A_f. If μ is large, what is the value of A_f?
(e) Find R_{in} and R_{out} assuming the loop gain is large.
(f) The "super" CG transistor can be utilized in the cascode configuration shown in Fig. P11.78(b), where V_G is a dc bias voltage. Replacing Q_1 by its small-signal model, use the analogy of the resulting circuit to that in Fig. P11.78(a) to find I_o and R_{out}.

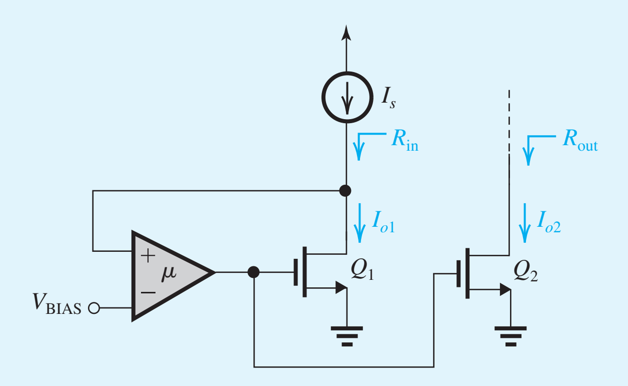


Figure P11.79

***11.79** Figure P11.79 shows an interesting and very useful application of feedback to improve the performance of the current mirror formed by Q_1 and Q_2. Rather than connecting the drain of Q_1 to the gate, as is the case in simple current mirrors, an amplifier of gain $+\mu$ is connected between the drain and the gate. Note that the feedback loop does not include transistor Q_2. The feedback loop ensures that the value of the gate-to-source voltage of Q_1 is such that I_{o1} equals I_s. This regulated V_{gs} is also applied to Q_2. Thus, if W/L of Q_2 is n times W/L of Q_1, $I_{o2} = nI_{o1} = nI_s$. This current tracking, however, is *not* regulated by the feedback loop.

(a) Show that the feedback is negative.

(b) If μ is very large and the input resistance of the amplifier μ is infinite, what dc voltage appears at the drain of Q_1? If Q_1 is to operate at an overdrive voltage of 0.2 V, what is the minimum value that V_{BIAS} must have?

(c) Replacing Q_1 by its small-signal model, find an expression for the small-signal input resistance R_{in} assuming finite gain but infinite input resistance for the amplifier μ. Note that here it is much easier to do the analysis directly than to use the feedback-analysis approach. For large μ, what does R_{in} become?

(d) What is the output resistance R_{out}?

***11.80** The circuit in Fig. P11.80 is an implementation of a particular circuit building block known as **second-generation current convoyer** (CCII). It has three terminals besides ground: x, y, and z. The heart of the circuit is the feedback amplifier consisting of the differential amplifier μ and the complementary source follower (Q_N, Q_P). (Note that this feedback circuit is one we have encountered a number of times in this chapter, albeit with only one source-follower transistor.) In the following, assume that the differential amplifier has a very large gain μ and infinite differential input resistance. Also, let the two current mirrors have unity current-transfer ratios.

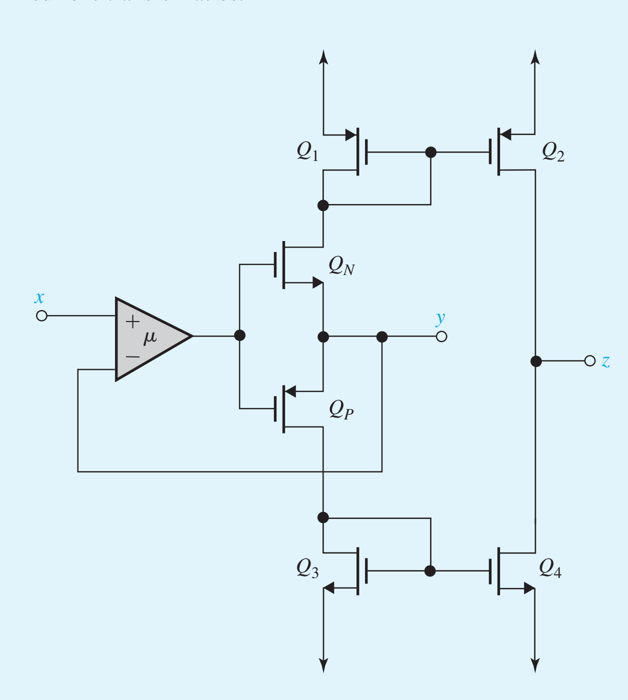


Figure P11.80

(a) If a resistance R is connected between y and ground, a voltage signal V_x is connected between x and ground, and z is short-circuited to ground. Find the current I_z through the short circuit. Show how this current is developed and its path for V_x positive and for V_x negative.

(b) If x is connected to ground, a current source I_y is connected to input terminal y, and z is connected to ground, what voltage appears at y and what is the input resistance seen by I_y? What is the current I_z that flows through the output short circuit? Also, explain the current flow through the circuit for I_y positive and for I_y negative.

(c) What is the output resistance at z?

SIM ***11.81** For the amplifier circuit in Fig. P11.81, assuming that V_s has a zero dc component, find the dc voltages at all nodes and the dc emitter currents of Q_1 and Q_2. Let the BJTs have $\beta = 100$. Use feedback analysis to find V_o/V_s and R_{in}. Let $V_{BE} = 0.7$ V.

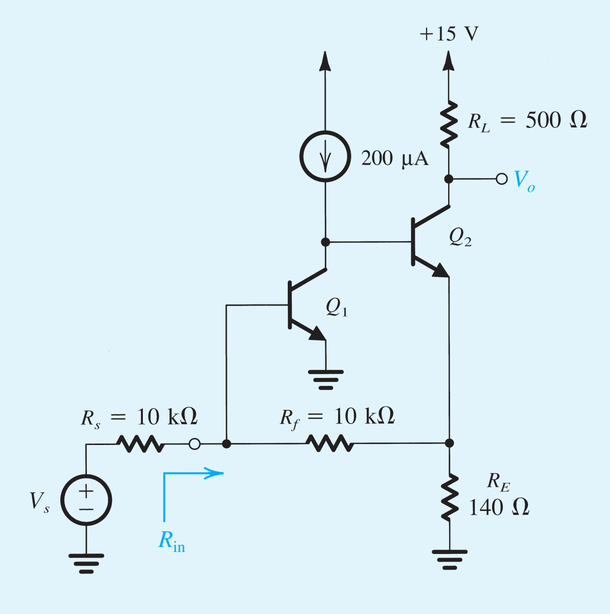


Figure P11.81

****11.82** Figure P11.82 shows a feedback amplifier utilizing the shunt–series topology. All transistors have $\beta = 100$ and $V_{BE} = 0.7$ V. Neglect r_o except in (f).

(a) Perform a dc analysis to find the dc emitter currents in Q_1 and Q_2 and hence determine their small-signal parameters.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

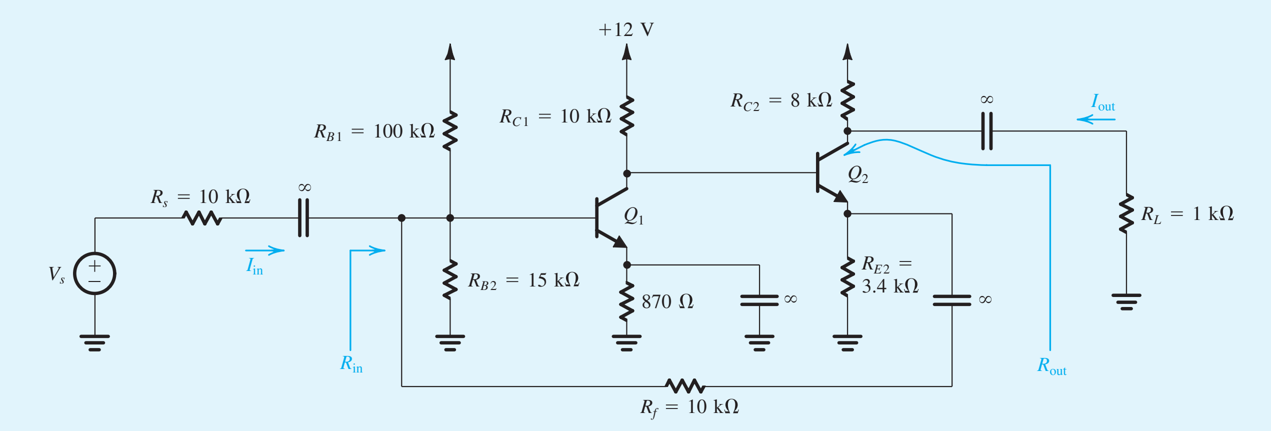


Figure P11.82

(b) Replacing the BJTs with their hybrid-π models, give the equivalent circuit of the feedback amplifier.

(c) Give the A circuit and determine A, R_i, and R_o. Note that R_o is the resistance determined by breaking the emitter loop of Q_2 and measuring the resistance between the terminals thus created.

(d) Find the β circuit and determine the value of β.

(e) Find $A\beta$, $1 + A\beta$, A_f, R_{if}, and R_{of}. Note that R_{of} represents the resistance that in effect appears in the emitter of Q_2 as a result of the feedback.

(f) Determine R_{in}, I_{out}/I_{in}, and R_{out}. To determine R_{out}, use $V_{A2} = 75$ V and recall that the maximum possible output resistance looking into the collector of a BJT is approximately βr_o, where β is the BJT's β (see Problem 11.55).

Section 11.7: The Stability Problem

11.83 An op amp designed to have a low-frequency gain of 10^5 and a high-frequency response dominated by a single pole at 100 rad/s acquires, through a manufacturing error, a pair of additional poles at 20,000 rad/s. At what frequency does the total phase shift reach 180°? At this frequency, for what value of β, assumed to be frequency independent, does the loop gain reach a value of unity? What is the corresponding value of closed-loop gain at low frequencies?

***11.84** For the situation described in Problem 11.83, sketch Nyquist plots for $\beta = 1.0$ and 10^{-3}. (Plot for $\omega = 0$ rad/s, 100 rad/s, 10^3 rad/s, 10^4 rad/s, 2×10^4 rad/s, and ∞ rad/s.)

11.85 An op amp having a low-frequency gain of 10^4 and a single-pole rolloff at 10^3 rad/s is connected in a negative-feedback loop via a feedback network having a transmission k and a two-pole rolloff at 10^3 rad/s. Find the value of k above which the closed-loop amplifier becomes unstable.

11.86 Consider a feedback amplifier for which the open-loop gain $A(s)$ is given by

$$A(s) = \frac{10{,}000}{(1 + s/10^4)(1 + s/10^5)^2}$$

If the feedback factor β is independent of frequency, find the frequency at which the phase shift is 180°, and find the critical value of β at which oscillation will commence.

Section 11.8: Effect of Feedback on the Amplifier Poles

11.87 A dc amplifier having a single-pole response with pole frequency 10 Hz and unity-gain frequency of 1 MHz is operated in a loop whose frequency-independent feedback factor is 0.1. Find the low-frequency gain, the 3-dB frequency, and the unity-gain frequency of the closed-loop amplifier. By what factor does the pole shift?

11.88 An amplifier has dc open-loop gain of 80 dB and a single pole with 100-Hz frequency. It is utilized to design a feedback amplifier with a 3-dB frequency of 10 kHz. What β is needed? What is the dc closed-loop gain realized? Give an expression for $A_f(s)$.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

***11.89** An amplifier having a low-frequency gain of 10^4 and poles at 10^4 Hz and 10^5 Hz is operated in a closed negative-feedback loop with a frequency-independent β.

(a) For what value of β do the closed-loop poles become coincident? At what frequency?
(b) What is the low-frequency, closed-loop gain corresponding to the situation in (a)? What is the value of the closed-loop gain at the frequency of the coincident poles?
(c) What is the value of Q corresponding to the situation in (a)?
(d) If β is increased by a factor of 10, what are the new pole locations? What is the corresponding pole Q?

D 11.90 A dc amplifier has an open-loop gain of 1000 and two poles, a dominant one at 1 kHz and a high-frequency one whose location can be controlled. It is required to connect this amplifier in a negative-feedback loop that provides a dc closed-loop gain of 10 and a maximally flat response. Find the required value of β and the frequency at which the second pole should be placed. What is the 3-dB frequency of the closed-loop amplifier?

11.91 Reconsider Example 11.11 with the circuit in Fig. 11.34, modified to incorporate a so-called tapered network, in which the components immediately adjacent to the amplifier input are raised in impedance to $C/10$ and $10R$. Find expressions for the resulting pole frequency ω_0 and Q factor. For what value of K do the poles coincide? For what value of K does the response become maximally flat? For what value of K does the circuit oscillate?

D 11.92 A feedback amplifier having a dc closed-loop gain of 10 and a maximally flat second-order response with a 3-dB frequency of 1 kHz is required. The open-loop amplifier utilizes a cascade of two identical amplifier stages, each having a single-pole frequency response. Find the values required for β, the 3-dB frequency, and the dc gain of each of the two amplifier stages. Give an expression for $A_f(s)$.

11.93 Three identical inverting amplifier stages, each characterized by a low-frequency gain K and a single-pole response with $f_{3dB} = 100$ kHz, are connected in a feedback loop with $\beta = 1$. What is the minimum value of K at which the circuit oscillates? What would the frequency of oscillation be?

Section 11.9: Stability Study Using Bode Plots

11.94 Reconsider Exercise 11.26 for the case of the op amp wired as a unity-gain buffer. At what frequency is $|A\beta| = 1$? What is the corresponding phase margin?

11.95 Reconsider Exercise 11.26 for the case of a manufacturing error introducing a second pole at 10^3 Hz. What is now the frequency for which $|A\beta| = 1$? What is the corresponding phase margin? For what values of β is the phase margin 45° or more?

11.96 For what phase margin does the gain peaking have a value of 5%? Of 10%? Of 0.1 dB? Of 1 dB? Of 3 dB? [*Hint:* Use the result in Eq. (11.82).]

11.97 An amplifier has a dc gain of 10^4 and poles at 10^5 Hz, 3.16×10^5 Hz, and 10^6 Hz. Find the value of β, and the corresponding closed-loop gain, for which a phase margin of 45° is obtained.

11.98 A two-pole amplifier for which $A_0 = 10^3$ and having poles at 1 MHz and 10 MHz is to be connected as a differentiator. On the basis of the rate-of-closure rule, what is the smallest differentiator time constant for which operation is stable? What are the corresponding gain and phase margins?

11.99 For the amplifier described by Fig. 11.37 and with frequency-independent feedback, what is the minimum closed-loop voltage gain that can be obtained for phase margins of 90° and 45°?

Section 11.10: Frequency Compensation

D 11.100 A multipole amplifier having a first pole at 1 MHz and a dc open-loop gain of 80 dB is to be compensated for closed-loop gains as low as unity by the introduction of a new dominant pole. At what frequency must the new pole be placed?

D 11.101 For the amplifier described in Problem 11.100, rather than introducing a new dominant pole we can use additional capacitance at the circuit node at which the pole is formed to reduce the frequency of the first pole. If the frequency of the second pole is 20 MHz and if it remains unchanged while additional capacitance is introduced as

mentioned, find the frequency to which the first pole must be lowered so that the resulting amplifier is stable for closed-loop gains as low as unity. By what factor is the capacitance at the controlling node increased?

11.102 For the amplifier whose $A(s)$ is depicted in Fig. 11.38, to what value must the first pole frequency be lowered to obtain stable performance for (a) $\beta = 0.001$ and (b) $\beta = 0.1$?

11.103 Contemplate the effects of pole splitting by considering Eqs. (11.89), (11.93), and (11.94) under the conditions that $R_1 \simeq R_2 = R$, $C_2 \simeq C_1/10 = C$, $C_f \gg C$, and $g_m = 100/R$, by calculating ω_{P1}, ω_{P2}, and ω'_{P1}, ω'_{P2}. Comment on the results.

D 11.104 An op amp with open-loop voltage gain of 10^5 and poles at 10^6 Hz, 10^7 Hz, and 10^8 Hz is to be compensated by the addition of a fourth dominant pole to operate stably with unity feedback ($\beta = 1$). What is the frequency of the required dominant pole? The compensation network is to consist of an *RC* low-pass network placed in the negative-feedback path of the op amp. The dc bias conditions are such that a 1-MΩ resistor can be tolerated in series with each of the negative and positive input terminals. What capacitor is required between the negative input and ground to implement the required fourth pole?

D *11.105 An op amp with an open-loop voltage gain of 80 dB and poles at 10^5 Hz, 10^6 Hz, and 2×10^6 Hz is to be compensated to be stable for unity β. Assume that the op amp incorporates an amplifier equivalent to that in Fig. 11.40, with $C_1 = 150$ pF, $C_2 = 5$ pF, and $g_m = 40$ mA/V, and that f_{P1} is caused by the input circuit and f_{P2} by the output circuit of this amplifier. Find the required value of the compensating Miller capacitance and the new frequency of the output pole.

SIM ****11.106** The op amp in the circuit of Fig. P11.106 has an open-loop gain of 10^5 and a single-pole rolloff with $\omega_{3dB} = 10$ rad/s.

(a) Sketch a Bode plot for the loop gain.
(b) Find the frequency at which $|A\beta| = 1$, and find the corresponding phase margin.
(c) Find the closed-loop transfer function, including its zero and poles. Sketch a pole-zero plot. Sketch the magnitude of the transfer function versus frequency, and label the important parameters on your sketch.

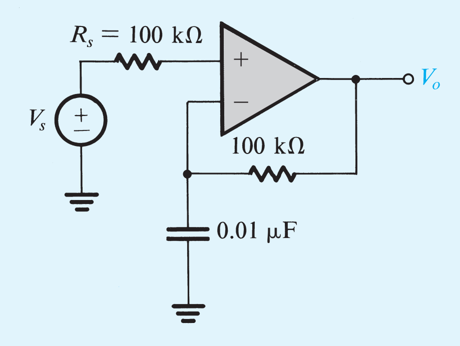


Figure P11.106

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

CHAPTER 12

Output Stages and Power Amplifiers

IN THIS CHAPTER YOU WILL LEARN

1. The classification of amplifier output stages on the basis of the fraction of the cycle of an input sine wave during which the transistor conducts.
2. Analysis and design of a variety of output-stage types ranging from the simple but power-inefficient emitter follower (class A) to the popular push–pull class AB circuit in both bipolar and CMOS technologies, and the power-efficient class D amplifier.
3. Useful and interesting circuit techniques employed in the design of power amplifiers.
4. The special structures and characteristics of bipolar and MOS power transistors.
5. Thermal considerations in the design and fabrication of high-output-power circuits.

Introduction

An important function of the output stage is to provide the amplifier with a low output resistance so that it can deliver the output signal to the load without loss of gain. Since the output stage is the final stage of the amplifier, it usually deals with relatively large signals. Thus the small-signal approximations and models either are not applicable or must be used with care. Nevertheless, linearity remains a very important requirement. In fact, a measure of goodness of the output stage is the amount of **total harmonic distortion** (THD) it introduces. This is the rms value of the harmonic components of the output signal, excluding the fundamental, expressed as a percentage of the rms value of the fundamental. A high-fidelity audio power amplifier features a THD in the order of a fraction of a percent.

The most challenging requirement in the design of an output stage is for it to deliver the required amount of power to the load in an *efficient* manner. This implies that the power *dissipated* in the output-stage transistors must be as low as possible. This requirement stems mainly from the fact that the power dissipated in a transistor raises its internal junction temperature (the temperature of the silicon die), and there is a maximum temperature (in the range of 150°C to 200°C for silicon devices) above which the transistor is destroyed. A high power-conversion efficiency is also required to prolong the battery life of portable electronics, to permit a smaller, lower-cost power supply, or to obviate the need for cooling fans.

We begin this chapter with a study of the various output-stage configurations employed in amplifiers that handle both low and high power. In this context, "high power" generally means greater than 1 W. Examples include the transmitter of a cell phone, which is typically required to deliver 1 W of power to its antenna, and a stereo system that delivers hundreds of watts of audio power to its speakers.

A power amplifier is simply an amplifier with a high-power output stage. Examples of discrete- and integrated-circuit power amplifiers will be presented. Of particular interest is the class D amplifier; it combines analog and digital techniques to achieve very high power-conversion efficiencies.

The chapter concludes with a study of power BJTs and MOSFETs. These are the devices employed in the various circuits studied in this chapter.

12.1 Classification of Output Stages

Output stages are classified according to the collector-current waveform that results when an input signal is applied. Figure 12.1 illustrates the classification for the case of a sinusoidal input signal. The class A stage, whose associated waveform is shown in Fig. 12.1(a), is biased at a current I_C that is greater than the amplitude of the signal current, $\hat{I}_c$. Thus the transistor in a class A stage conducts for the entire cycle of the input signal; that is, the conduction angle is 360°. In contrast, the class B stage, whose associated waveform is shown in Fig. 12.1(b),

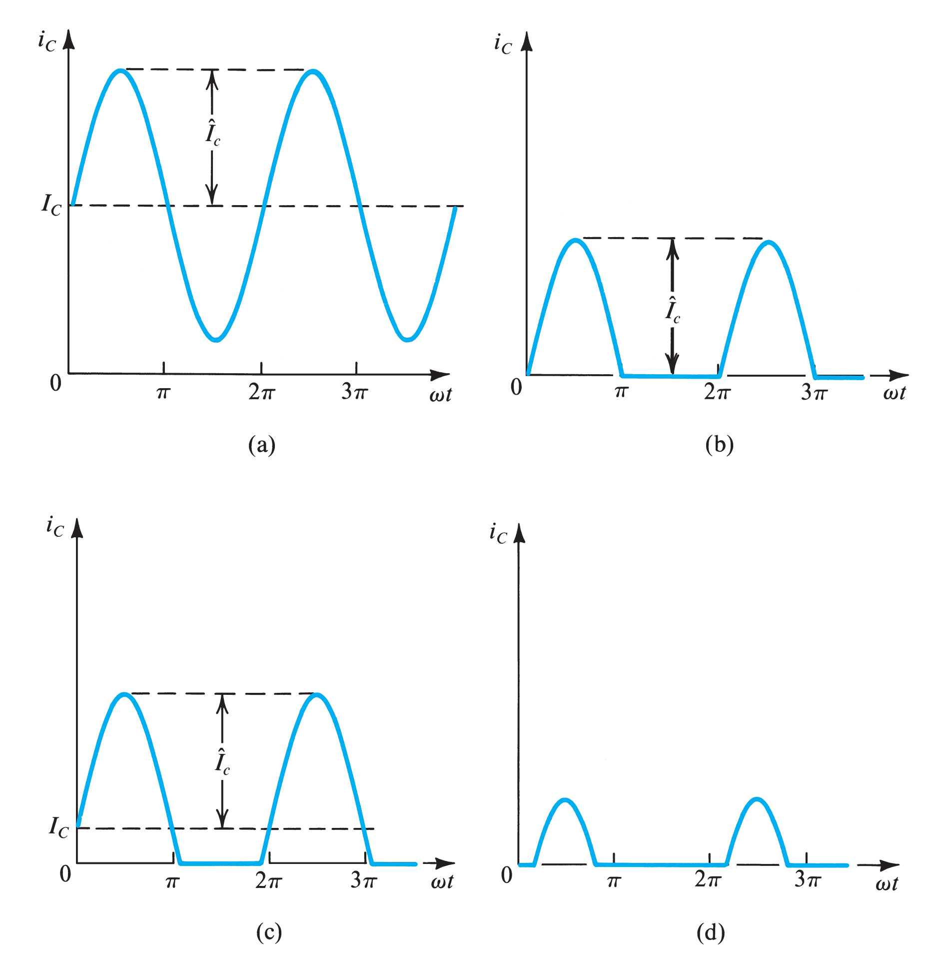


Figure 12.1 Collector-current waveforms for transistors operating in **(a)** class A, **(b)** class B, **(c)** class AB, and **(d)** class C amplifier stages.

is biased at zero dc current. Thus a transistor in a class B stage conducts for only half the cycle of the input sine wave, resulting in a conduction angle of 180°. As will be seen later, the negative halves of the sinusoid will be supplied by another transistor that also operates in the class B mode and conducts during the alternate half-cycles.

An intermediate class between A and B, appropriately named class AB, involves biasing the transistor at a nonzero dc current much smaller than the peak current of the sine-wave signal. As a result, the transistor conducts for an interval slightly greater than half a cycle, as illustrated in Fig. 12.1(c). The resulting conduction angle is greater than 180° but much less than 360°. The class AB stage has another transistor that conducts for an interval slightly greater than that of the negative half-cycle, and the currents from the two transistors are combined in the load. It follows that, during the intervals near the zero crossings of the input sinusoid, both transistors conduct.

Figure 12.1(d) shows the collector-current waveform for a transistor operated as a class C amplifier. Observe that the transistor conducts for an interval shorter than that of a half-cycle; that is, the conduction angle is less than 180°. The result is the periodically pulsating current waveform shown. To obtain a sinusoidal output voltage, this current is passed through a parallel *LC* circuit, tuned to the frequency of the input sinusoid. The tuned circuit acts as a bandpass filter (Chapter 17) and provides an output voltage proportional to the amplitude of the fundamental component in the Fourier-series representation of the current waveform.

Class A, AB, and B amplifiers are studied in this chapter. They are employed as output stages of op amps and audio power amplifiers. In the latter application, class AB is the preferred choice, for reasons that will be explained in the sections to follow. A variation on the class AB, called the class G/H amplifier, utilizes two pairs of power supplies (e.g., the regular ±30-V supply and a higher-voltage supply of ±70 V). The higher-voltage supply is called upon only occasionally—for instance, to provide a short burst of high output power for a drum roll. Since the high-voltage supply operates infrequently, it can be of a low-cost design. Also, this is a more power-efficient arrangement than would be obtained if a class AB circuit were used and operated continuously from the higher-voltage supply.

Class C amplifiers are usually employed for radio-frequency (RF) power amplification (required, e.g., in mobile phones and radio and TV transmitters). The design of class C amplifiers is a rather specialized topic and is not included in this book. However, we should point out that the tuned-resonator oscillator circuits described in Chapter 18 operate inherently in the class C mode.

Another important type of output stage is the class D switching amplifier. It utilizes the power transistors as on-off switches and thus achieves much higher power efficiency than that obtained in the other amplifier classes. The class D amplifier will be studied briefly in Section 12.9. Although the BJT in Fig. 12.1 has been used to illustrate the definition of the various output-stage classes, the same classification applies to output stages implemented with MOSFETs. Furthermore, the classification above extends to amplifier stages other than those used at the output. In this regard, all the common-emitter, common-base, and common-collector amplifiers (and their FET counterparts) studied in earlier chapters fall into the class A category.

12.2 Class A Output Stage

Because of its low output resistance, the emitter follower is the most popular class A output stage. We have already studied the emitter follower in Chapter 7; in the following we consider its large-signal operation.

12.2.1 Transfer Characteristic

Figure 12.2 shows an emitter follower Q_1 biased with a constant current I supplied by transistor Q_2. Since the emitter current $i_{E1} = I + i_L$, the bias current I must be greater than the largest negative load current; otherwise, Q_1 cuts off and class A operation will no longer be maintained.

The transfer characteristic of the emitter follower of Fig. 12.2 is described by

$$v_O = v_I - v_{BE1} \tag{12.1}$$

where v_{BE1} depends on the emitter current i_{E1} and thus on the load current i_L. If we neglect the relatively small changes in v_{BE1} (60 mV for every factor-of-10 change in emitter current), the linear transfer curve shown in Fig. 12.3 results. As indicated, the positive limit of the linear region is determined by the saturation of Q_1; thus

$$v_{O\max} = V_{CC} - V_{CE1\text{sat}} \tag{12.2}$$

In the negative direction, depending on the values of I and R_L, the limit of the linear region is determined either by Q_1 turning off,

$$v_{O\min} = -IR_L \tag{12.3}$$

or by Q_2 saturating,

$$v_{O\min} = -V_{CC} + V_{CE2\text{sat}} \tag{12.4}$$

The absolutely lowest (most negative) output voltage is that given by Eq. (12.4) and is achieved provided the bias current I is greater than the magnitude of the corresponding load current,

$$I \geq \frac{|-V_{CC} + V_{CE2\text{sat}}|}{R_L} \tag{12.5}$$

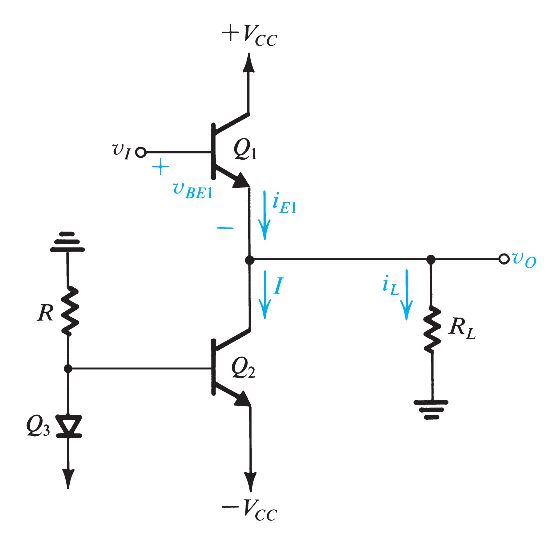


Figure 12.2 An emitter follower (Q_1) biased with a constant current I supplied by transistor Q_2.

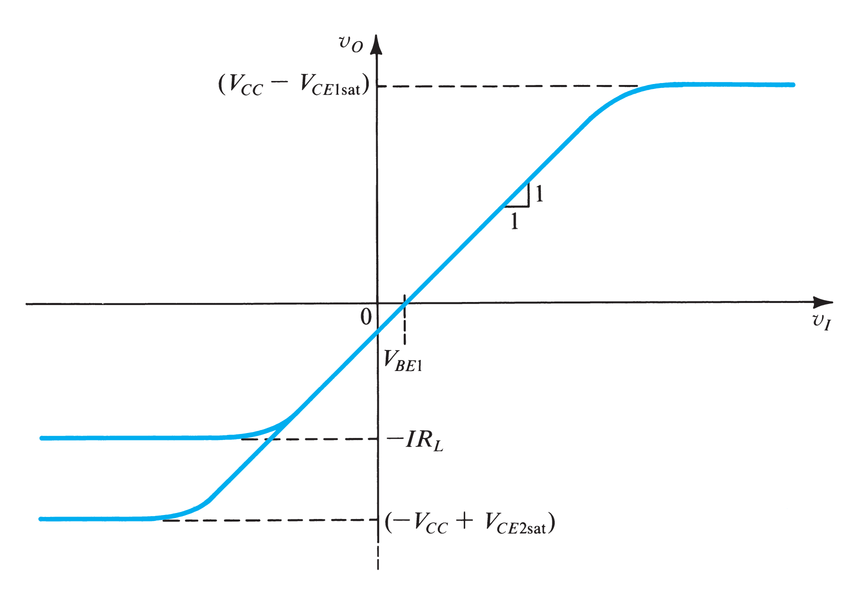


Figure 12.3 Transfer characteristic of the emitter follower in Fig. 12.2. This linear characteristic is obtained by neglecting the change in v_{BE1} with i_L. The maximum positive output is determined by the saturation of Q_1. In the negative direction, the limit of the linear region is determined either by Q_1 turning off or by Q_2 saturating, depending on the values of I and R_L.

EXERCISES

D12.1 For the emitter follower in Fig. 12.2, $V_{CC} = 15$ V, $V_{CEsat} = 0.2$ V, $V_{BE} = 0.7$ V and constant, and β is very high. Find the value of R that will establish a bias current sufficiently large to allow the largest possible output signal swing for $R_L = 1$ kΩ. Determine the resulting output signal swing and the minimum and maximum emitter currents for Q_1.
Ans. 0.97 kΩ; –14.8 V to +14.8 V; 0 to 29.6 mA

12.2 For the emitter follower of Exercise 12.1, in which $I = 14.8$ mA and $R_L = 1$ kΩ, consider the case in which v_O is limited to the range –10 V to +10 V. Let Q_1 have $v_{BE} = 0.6$ V at $i_C = 1$ mA, and assume $\alpha \simeq 1$. Find v_I corresponding to $v_O = -10$ V, 0 V, and +10 V. At each of these points, use small-signal analysis to determine the voltage gain v_o/v_i. Note that the incremental voltage gain gives the slope of the v_O-versus-v_I characteristic.
Ans. –9.36 V, 0.67 V, 10.68 V; 0.995 V/V, 0.998 V/V, 0.999 V/V

12.2.2 Signal Waveforms

Consider the operation of the emitter-follower circuit of Fig. 12.2 for sine-wave input. Neglecting V_{CEsat}, we see that if the bias current I is properly selected, the output voltage can swing from $-V_{CC}$ to $+V_{CC}$ with the quiescent value being zero, as shown in Fig. 12.4(a).

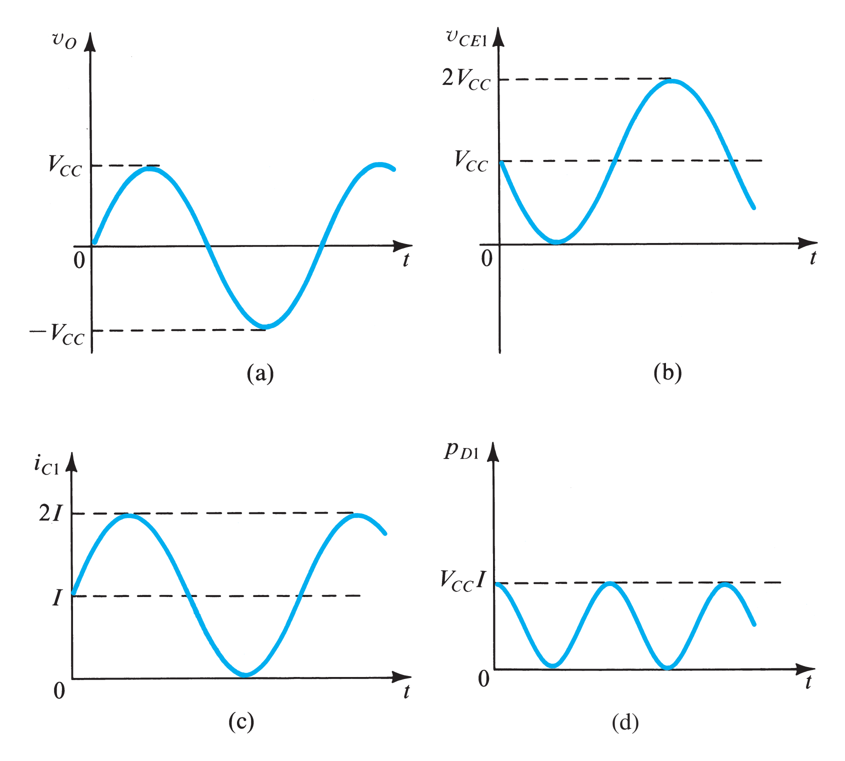


Figure 12.4 Maximum signal waveforms in the class A output stage of Fig. 12.2 under the condition $I = V_{CC}/R_L$ or, equivalently, $R_L = V_{CC}/I$. Note that the transistor saturation voltages have been neglected.

Figure 12.4(b) shows the corresponding waveform of $v_{CE1} = V_{CC} - v_O$. Now, assuming that the bias current I is selected to allow a maximum negative load current of V_{CC}/R_L, that is,

$$I = V_{CC}/R_L$$

the collector current of Q_1 will have the waveform shown in Fig. 12.4(c). Finally, Fig. 12.4(d) shows the waveform of the **instantaneous power dissipation** in Q_1,

$$p_{D1} \equiv v_{CE1} i_{C1} \tag{12.6}$$

12.2.3 Power Dissipation

Figure 12.4(d) indicates that the maximum instantaneous power dissipation in Q_1 is $V_{CC}I$. This is equal to the power dissipation in Q_1 with no input signal applied, that is, the quiescent power dissipation. Thus the emitter-follower transistor dissipates the largest amount of power when $v_O = 0$. Since this condition (no input signal) can easily prevail for prolonged periods of time, transistor Q_1 must be able to withstand a continuous power dissipation of $V_{CC}I$.

The power dissipation in Q_1 depends on the value of R_L. Consider the extreme case of an output open circuit, that is, $R_L = \infty$. In this case, $i_{C1} = I$ is constant and the instantaneous power dissipation in Q_1 will depend on the instantaneous value of v_O. The maximum power dissipation will occur when $v_O = -V_{CC}$, for in this case v_{CE1} is a maximum of $2V_{CC}$ and $p_{D1} = 2V_{CC}I$. This condition, however, would not normally persist for a prolonged interval, so the design need not be that conservative. Observe that with an open-circuit load, the

average power dissipation in Q_1 is $V_{CC}I$. A far more dangerous situation occurs at the other extreme of R_L—specifically, $R_L = 0$. In the event of an output short circuit, a positive input voltage would theoretically result in an infinite load current. In practice, a very large current may flow through Q_1, and if the short-circuit condition persists, the resulting large power dissipation in Q_1 can raise its junction temperature beyond the maximum allowed, causing permanent damage. To guard against such a situation, output stages are usually equipped with **short-circuit protection**, as will be explained later.

The power dissipation in Q_2 also must be taken into account in designing an emitter-follower output stage. Since Q_2 conducts a constant current I, and the maximum value of v_{CE2} is $2V_{CC}$, the maximum instantaneous power dissipation in Q_2 is $2V_{CC}I$. This maximum, however, occurs when $v_O = V_{CC}$, a condition that would not normally prevail for a prolonged period of time. A more significant quantity for design purposes is the average power dissipation in Q_2, which is $V_{CC}I$.

Example 12.1

Consider the emitter follower in Fig. 12.2 with $V_{CC} = 10$ V, $I = 100$ mA, and $R_L = 100\,\Omega$.

(a) Find the power dissipated in Q_1 and Q_2 under quiescent conditions $(v_O = 0)$.
(b) For a sinusoidal output voltage of maximum possible amplitude (neglecting V_{CEsat}), find the average power dissipation in Q_1 and Q_2. Also find the load power.

Solution

(a) Under quiescent conditions $v_O = 0$, and each of Q_1 and Q_2 conducts a current $I = 100$ mA $= 0.1$ A and has a voltage $V_{CE} = V_{CC} = 10$ V, thus

$$P_{D1} = P_{D2} = V_{CC}I = 10 \times 0.1 = 1 \text{ W}$$

(b) For a sinusoidal output voltage of maximum possible amplitude (i.e., 10-V peak), the instantaneous power dissipation in Q_1 will be as shown in Fig. 12.4(d). Thus the average power dissipation in Q_1 will be

$$P_{D1} = \frac{1}{2}V_{CC}I = \frac{1}{2} \times 10 \times 0.1 = 0.5 \text{ W}$$

For Q_2, the current is constant at $I = 0.1$ A and the voltage at the collector will have an average value of 0 V. Thus the average voltage across Q_2 will be V_{CC} and the average dissipation will be

$$P_{D2} = I \times v_{CE}\big|_{\text{average}}$$
$$= I \times V_{CC} = 0.1 \times 10 = 1 \text{ W}$$

Finally, the power delivered to the load can be found from

$$P_L = \frac{V_{orms}^2}{R_L}$$
$$= \frac{(10/\sqrt{2})^2}{100} = 0.5 \text{ W}$$

12.2.4 Power-Conversion Efficiency

The power-conversion efficiency of an output stage is defined as

$$\eta \equiv \frac{\text{Load power}(P_L)}{\text{Supply power}(P_S)} \tag{12.7}$$

For the emitter follower of Fig. 12.2, assuming that the output voltage is a sinusoid with the peak value $\hat{V}_o$, the average load power will be

$$P_L = \frac{(\hat{V}_o/\sqrt{2})^2}{R_L} = \frac{1}{2}\frac{\hat{V}_o^2}{R_L} \tag{12.8}$$

Since the current in Q_2 is constant (I), the power drawn from the negative supply[1] is $V_{CC}I$. The *average* current in Q_1 is equal to I, and thus the average power drawn from the positive supply is $V_{CC}I$. Thus the total average supply power is

$$P_S = 2V_{CC}I \tag{12.9}$$

Equations (12.8) and (12.9) can be combined to yield

$$\eta = \frac{1}{4}\frac{\hat{V}_o^2}{IR_L V_{CC}}$$

$$= \frac{1}{4}\left(\frac{\hat{V}_o}{IR_L}\right)\left(\frac{\hat{V}_o}{V_{CC}}\right) \tag{12.10}$$

Since $\hat{V}_o \leq V_{CC}$ and $\hat{V}_o \leq IR_L$, maximum efficiency is obtained when

$$\hat{V}_o = V_{CC} = IR_L \tag{12.11}$$

The maximum efficiency attainable is 25%. Because this is a rather low figure, the class A output stage is rarely used in high-power applications (>1 W). Note also that in practice the output voltage swing is limited to lower values to avoid transistor saturation and associated nonlinear distortion. Thus the efficiency achieved in practice is usually in the 10% to 20% range.

EXERCISE

12.3 For the emitter follower of Fig. 12.2, let $V_{CC} = 10$ V, $I = 100$ mA, and $R_L = 100\ \Omega$. If the output voltage is an 8-V-peak sinusoid, find the following: (a) the power delivered to the load; (b) the average power drawn from the supplies; (c) the power-conversion efficiency. Ignore the loss in Q_3 and R.
Ans. 0.32 W; 2 W; 16%

[1] This does *not* include the power drawn by the biasing resistor R and the diode-connected transistor Q_3.

12.3 Class B Output Stage

Figure 12.5 shows a class B output stage. It consists of a complementary pair of transistors (an *npn* and a *pnp*) connected in such a way that both cannot conduct simultaneously.

12.3.1 Circuit Operation

When the input voltage v_I is zero, both transistors are cut off and the output voltage v_O is zero. As v_I goes positive and exceeds about 0.5 V, Q_N conducts and operates as an emitter follower. In this case v_O follows v_I (i.e., $v_O = v_I - v_{BEN}$) and Q_N supplies the load current. Meanwhile, the emitter–base junction of Q_P will be reverse biased by the V_{BE} of Q_N, which is approximately 0.7 V. Thus Q_P will be cut off.

If the input goes negative by more than about 0.5 V, Q_P turns on and acts as an emitter follower. Again v_O follows v_I (i.e., $v_O = v_I + v_{EBP}$), but in this case Q_P supplies the load current (in the direction opposite to that of i_L, since v_O will be negative), and Q_N will be cut off.

We conclude that the transistors in the class B stage of Fig. 12.5 are biased at zero current and conduct only when the input signal is present. The circuit operates in a **push–pull** fashion: Q_N *pushes* (sources) current into the load when v_I is positive, and Q_P *pulls* (sinks) current from the load when v_I is negative.

12.3.2 Transfer Characteristic

A sketch of the transfer characteristic of the class B stage is shown in Fig. 12.6. Note that there exists a range of v_I centered around zero where both transistors are cut off and v_O is zero. This **dead band** results in the **crossover distortion** illustrated in Fig. 12.7 for the case of an input sine wave. The effect of crossover distortion will be most pronounced when the amplitude of the input signal is small. Crossover distortion in audio power amplifiers gives rise to unpleasant sounds.

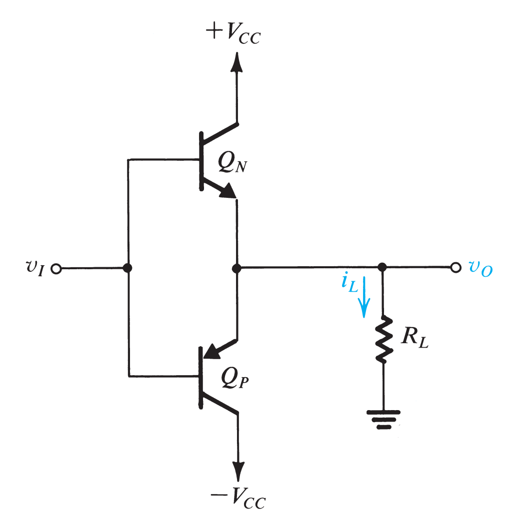


Figure 12.5 A class B output stage.

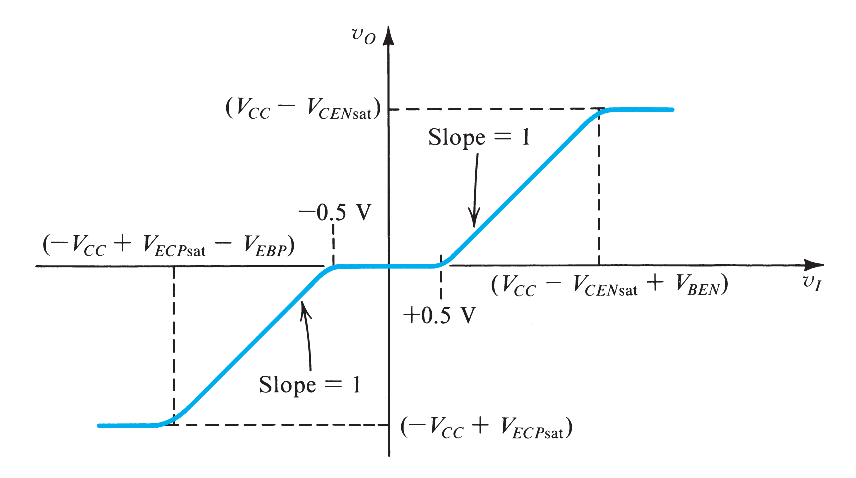


Figure 12.6 Transfer characteristic for the class B output stage in Fig. 12.5.

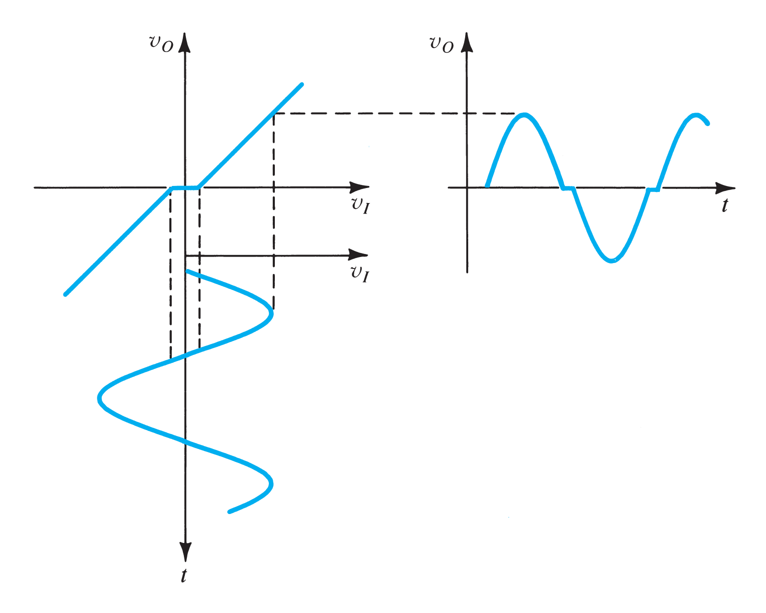


Figure 12.7 Illustrating how the dead band in the class B transfer characteristic results in crossover distortion.

12.3.3 Power-Conversion Efficiency

To calculate the power-conversion efficiency, η, of the class B stage, we neglect the crossover distortion and consider the case of an output sinusoid of peak amplitude $\hat{V}_o$. The average load power will be

$$P_L = \frac{1}{2}\frac{\hat{V}_o^2}{R_L} \tag{12.12}$$

The current drawn from each supply will consist of half-sine waves of peak amplitude $(\hat{V}_o/R_L)$. Thus the average current drawn from each of the two power supplies will be $\hat{V}_o/\pi R_L$. It follows that the average power drawn from each of the two power supplies will be the same,

$$P_{S+} = P_{S-} = \frac{1}{\pi}\frac{\hat{V}_o}{R_L}V_{CC} \tag{12.13}$$

and the total supply power will be

$$P_S = \frac{2}{\pi}\frac{\hat{V}_o}{R_L}V_{CC} \tag{12.14}$$

Thus the efficiency will be given by

$$\eta = \left(\frac{1}{2}\frac{\hat{V}_o^2}{R_L}\right) \Bigg/ \left(\frac{2}{\pi}\frac{\hat{V}_o}{R_L}V_{CC}\right) = \frac{\pi}{4}\frac{\hat{V}_o}{V_{CC}} \tag{12.15}$$

It follows that the maximum efficiency is obtained when $\hat{V}_o$ is at its maximum. This maximum is limited by the saturation of Q_N and Q_P to $V_{CC} - V_{CE\text{sat}} \simeq V_{CC}$. At this value of peak output voltage, the power-conversion efficiency is

$$\eta_{\max} = \frac{\pi}{4} = 78.5\% \tag{12.16}$$

This value is much larger than that obtained in the class A stage (25%). Finally, we note that the maximum average power available from a class B output stage is obtained by substituting $\hat{V}_o = V_{CC}$ in Eq. (12.12),

$$P_{L\max} = \frac{1}{2}\frac{V_{CC}^2}{R_L} \tag{12.17}$$

12.3.4 Power Dissipation

Unlike the class A stage, which dissipates maximum power under quiescent conditions ($v_O = 0$), the quiescent power dissipation of the class B stage is zero. When an input signal is applied, the *average* power dissipated in the class B stage is given by

$$P_D = P_S - P_L \tag{12.18}$$

Substituting for P_S from Eq. (12.14) and for P_L from Eq. (12.12) results in

$$P_D = \frac{2}{\pi}\frac{\hat{V}_o}{R_L}V_{CC} - \frac{1}{2}\frac{\hat{V}_o^2}{R_L} \tag{12.19}$$

From symmetry we see that half of P_D is dissipated in Q_N and the other half in Q_P. Thus Q_N and Q_P must be capable of safely dissipating $\frac{1}{2}P_D$ watts. Since P_D depends on $\hat{V}_o$, we must find the worst-case power dissipation, $P_{D\max}$. Differentiating Eq. (12.19) with respect to $\hat{V}_o$ and equating the derivative to zero gives the value of $\hat{V}_o$ that results in maximum average power dissipation as

$$\hat{V}_o\big|_{P_{D\max}} = \frac{2}{\pi}V_{CC} \tag{12.20}$$

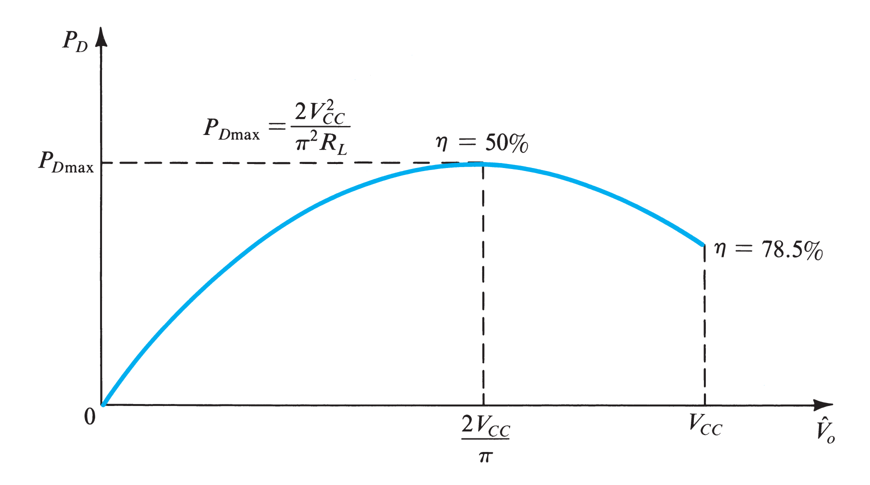


Figure 12.8 Power dissipation of the class B output stage versus amplitude of the output sinusoid.

Substituting this value in Eq. (12.19) gives

$$P_{D\max} = \frac{2V_{CC}^2}{\pi^2 R_L} \tag{12.21}$$

Thus,

$$P_{DN\max} = P_{DP\max} = \frac{V_{CC}^2}{\pi^2 R_L} \tag{12.22}$$

At the point of maximum power dissipation, the efficiency can be evaluated by substituting for $\hat{V}_o$ from Eq. (12.20) into Eq. (12.15); hence, $\eta = 50\%$.

Figure 12.8 shows a sketch of P_D (Eq. 12.19) versus the peak output voltage $\hat{V}_o$. Curves such as this are usually given on the data sheets of IC power amplifiers. [Usually, however, P_D is plotted versus P_L, as $P_L = \frac{1}{2}(\hat{V}_o^2/R_L)$ rather than $\hat{V}_o$.] An interesting observation follows from Fig. 12.8: Increasing $\hat{V}_o$ beyond $2V_{CC}/\pi$ *decreases* the power dissipated in the class B stage while increasing the load power. The price paid is an increase in nonlinear distortion as a result of approaching the saturation region of operation of Q_N and Q_P. Transistor saturation flattens the peaks of the output sine waveform. Unfortunately, this type of distortion cannot be significantly reduced by the application of negative feedback (see Section 11.2), and thus transistor saturation should be avoided in applications requiring low THD.

Example 12.2

It is required to design a class B output stage to deliver an average power of 20 W to an 8-Ω load. The power supply is to be selected such that V_{CC} is about 5 V greater than the peak output voltage. This avoids transistor saturation and the associated nonlinear distortion, and allows for including short-circuit protection circuitry. (The latter will be discussed in Section 12.6.) Determine the supply voltage required,

the peak current drawn from each supply, the total supply power, and the power-conversion efficiency. Also determine the maximum power that each transistor must be able to dissipate safely.

Solution

Since

$$P_L = \frac{1}{2}\frac{\hat{V}_o^2}{R_L}$$

then

$$\hat{V}_o = \sqrt{2P_L R_L}$$
$$= \sqrt{2 \times 20 \times 8} = 17.9 \text{ V}$$

Therefore we select $V_{CC} = 23$ V.

The peak current drawn from each supply is

$$\hat{I}_o = \frac{\hat{V}_o}{R_L} = \frac{17.9}{8} = 2.24 \text{ A}$$

Since each supply provides a current waveform of half-sinusoids, the average current drawn from each supply will be $\hat{I}_o/\pi$. Thus the average power drawn from each supply is

$$P_{S+} = P_{S-} = \frac{1}{\pi} \times 2.24 \times 23 = 16.4 \text{ W}$$

for a total supply power of 32.8 W. The power-conversion efficiency is

$$\eta = \frac{P_L}{P_S} = \frac{20}{32.8} \times 100 = 61\%$$

The maximum power dissipated in each transistor is given by Eq. (12.22); thus,

$$P_{DN\max} = P_{DP\max} = \frac{V_{CC}^2}{\pi^2 R_L}$$
$$= \frac{(23)^2}{\pi^2 \times 8} = 6.7 \text{ W}$$

12.3.5 Reducing Crossover Distortion

The crossover distortion of a class B output stage can be reduced substantially by employing a high-gain op amp and overall negative feedback, as shown in Fig. 12.9. The ± 0.7-V dead band is reduced to $\pm 0.7/A_0$ volt, where A_0 is the dc gain of the op amp. Nevertheless, the slew-rate limitation of the op amp will cause the alternate turning on and off of the output transistors to be noticeable, especially at high frequencies. A more practical method for reducing and almost eliminating crossover distortion is found in the class AB stage, which will be studied in the next section.

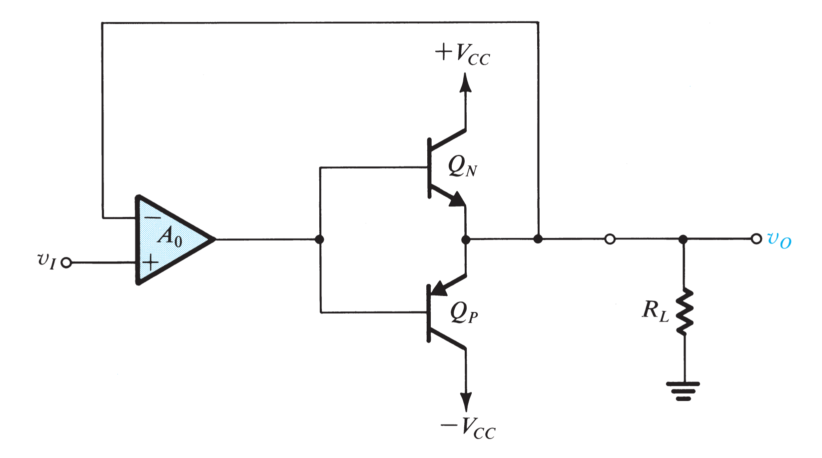


Figure 12.9 Class B circuit with an op amp connected in a negative-feedback loop to reduce crossover distortion.

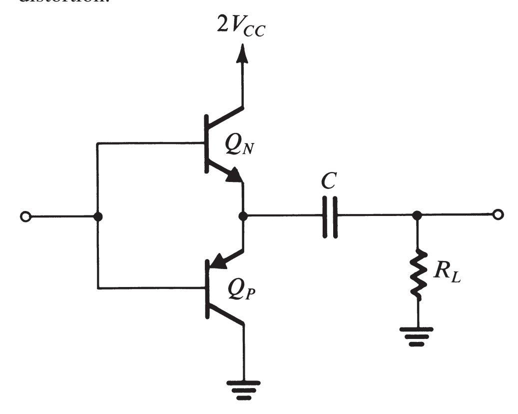


Figure 12.10 Class B output stage operated with a single power supply.

12.3.6 Single-Supply Operation

The class B stage can be operated from a single power supply, in which case the load is capacitively coupled, as shown in Fig. 12.10. Note that to make the formulas derived in Section 12.3.4 directly applicable, the single power supply is denoted $2V_{CC}$.

EXERCISE

12.4 For the class B output stage of Fig. 12.5, let $V_{CC} = 6$ V and $R_L = 4\ \Omega$. If the output is a sinusoid with 4.5-V peak amplitude, find (a) the output power; (b) the average power drawn from each supply; (c) the power efficiency obtained at this output voltage; (d) the peak currents supplied by v_I, assuming that $\beta_N = \beta_P = 50$; and (e) the maximum power that each transistor must be capable of dissipating safely.

Ans. (a) 2.53 W; (b) 2.15 W; (c) 59%; (d) 22.1 mA; (e) 0.91 W

12.4 Class AB Output Stage

Crossover distortion can be virtually eliminated by biasing the complementary output transistors at a small nonzero current. The result is the class AB output stage shown in Fig. 12.11. A bias voltage V_{BB} is applied between the bases of Q_N and Q_P. For $v_I = 0$, $v_O = 0$, and a voltage $V_{BB}/2$ appears across the base–emitter junction of each of Q_N and Q_P. Assuming matched devices,

$$i_N = i_P = I_Q = I_S e^{V_{BB}/2V_T} \tag{12.23}$$

The value of V_{BB} is selected to yield the required quiescent current I_Q.

12.4.1 Circuit Operation

When v_I goes positive by a certain amount, the voltage at the base of Q_N increases by the same amount, and the output becomes positive at an almost equal value,

$$v_O = v_I + \frac{V_{BB}}{2} - v_{BEN} \tag{12.24}$$

The positive v_O causes a current i_L to flow through R_L, and thus i_N must increase; that is,

$$i_N = i_P + i_L \tag{12.25}$$

The increase in i_N will be accompanied by a corresponding increase in v_{BEN} (above the quiescent value of $V_{BB}/2$). However, since the voltage between the two bases remains constant

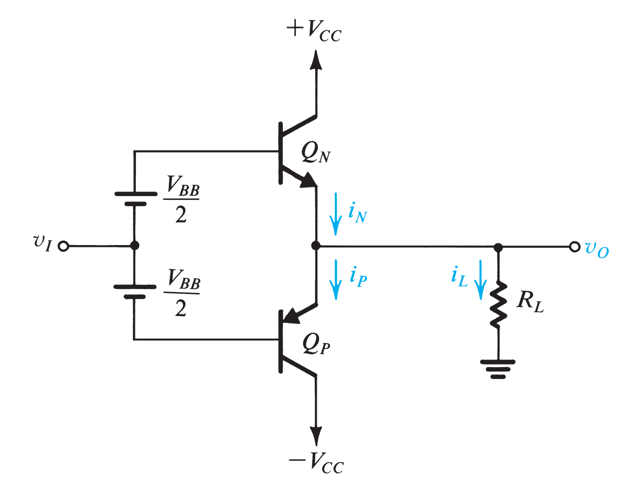


Figure 12.11 Class AB output stage. A bias voltage V_{BB} is applied between the bases of Q_N and Q_P, giving rise to a bias current I_Q given by Eq. (12.23). Thus, for small v_I, both transistors conduct and crossover distortion is almost completely eliminated.

at V_{BB}, the increase in v_{BEN} will result in an equal decrease in v_{EBP} and hence in i_P. The relationship between i_N and i_P can be derived as follows:

$$v_{BEN} + v_{EBP} = V_{BB}$$

$$V_T \ln \frac{i_N}{I_S} + V_T \ln \frac{i_P}{I_S} = 2V_T \ln \frac{I_Q}{I_S}$$

$$i_N i_P = I_Q^2 \tag{12.26}$$

Thus, as i_N increases, i_P decreases by the same ratio while the product remains constant. Equations (12.25) and (12.26) can be combined to yield i_N for a given i_L as the solution to the quadratic equation

$$i_N^2 - i_L i_N - I_Q^2 = 0 \tag{12.27}$$

From the equations above, we can see that for positive output voltages, the load current is supplied by Q_N, which acts as the output emitter follower. Meanwhile, Q_P will be conducting a current that decreases as v_O increases; for large v_O the current in Q_P can be ignored altogether.

For negative input voltages the opposite occurs: The load current will be supplied by Q_P, which acts as the output emitter follower, while Q_N conducts a current that gets smaller as v_I becomes more negative. Equation (12.26), relating i_N and i_P, holds for negative inputs as well.

We conclude that the class AB stage operates in much the same manner as the class B circuit, with one important exception: For small v_I, both transistors conduct, and as v_I is increased or decreased, one of the two transistors takes over the operation. Since the transition is a smooth one, crossover distortion will be almost totally eliminated. Figure 12.12 shows the transfer characteristic of the class AB stage.

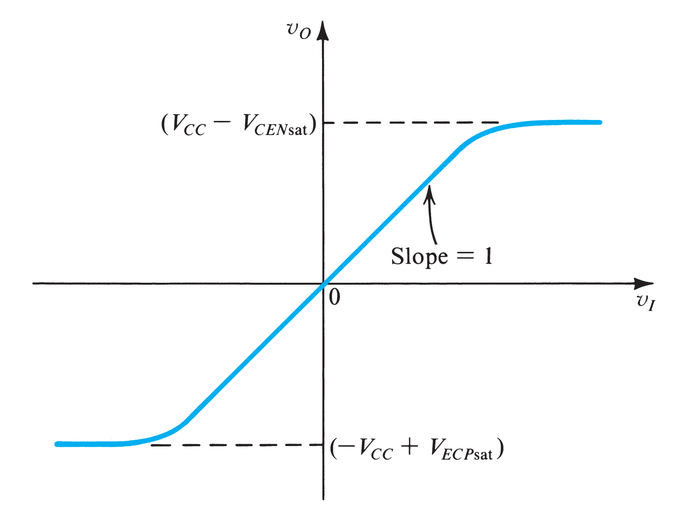


Figure 12.12 Transfer characteristic of the class AB stage in Fig. 12.11.

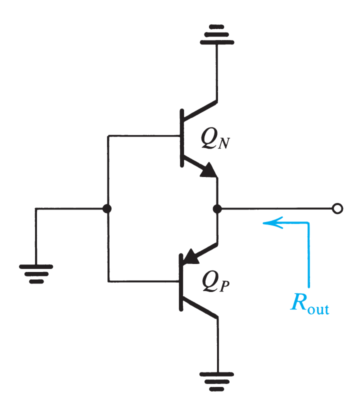


Figure 12.13 Determining the small-signal output resistance of the class AB circuit of Fig. 12.11.

The power relationships in the class AB stage are almost identical to those derived for the class B circuit in Section 12.3. The only difference is that under quiescent conditions the class AB circuit dissipates a power of $V_{CC}I_Q$ per transistor. Since I_Q is usually much smaller than the peak load current, the quiescent power dissipation is usually small. Nevertheless, it can be taken into account easily. Specifically, we can simply add the quiescent dissipation per transistor to its maximum power dissipation with an input signal applied, to obtain the total power dissipation that the transistor must be able to handle safely.

12.4.2 Output Resistance

If we assume that the source supplying v_I is ideal, then the output resistance of the class AB stage can be determined from the circuit in Fig. 12.13 as

$$R_{out} = r_{eN} \parallel r_{eP} \tag{12.28}$$

where r_{eN} and r_{eP} are the small-signal emitter resistances of Q_N and Q_P, respectively. At a given input voltage, the currents i_N and i_P can be determined, and r_{eN} and r_{eP} are given by

$$r_{eN} = \frac{V_T}{i_N} \tag{12.29}$$

$$r_{eP} = \frac{V_T}{i_P} \tag{12.30}$$

Thus,

$$R_{out} = \frac{V_T}{i_N} \bigg\| \frac{V_T}{i_P} = \frac{V_T}{i_P + i_N} \tag{12.31}$$

Since as i_N increases, i_P decreases, and vice versa, the output resistance remains approximately constant in the region around $v_I = 0$. This, in effect, is the reason for the virtual absence of crossover distortion. At larger load currents, either i_N or i_P will be significant, and R_{out} decreases as the load current increases.

Example 12.3

In this example we explore the details of the transfer characteristic, v_O versus v_I, of the class AB circuit in Fig. 12.11. For this purpose let $V_{CC} = 15$ V, $I_Q = 2$ mA, and $R_L = 100\ \Omega$. Assume that Q_N and Q_P are matched and have $I_S = 10^{-13}$ A. First, determine the required value of the bias voltage V_{BB}. Then, find the transfer characteristic for v_O in the range -10 V to $+10$ V.

Solution

To determine the required value of V_{BB} we use Eq. (12.23) with $I_Q = 2$ mA and $I_S = 10^{-13}$ A. Thus,

$$V_{BB} = 2V_T \ln\left(I_Q/I_S\right)$$
$$= 2 \times 0.025 \ln\left(2 \times 10^{-3}/10^{-13}\right) = 1.186 \text{ V}$$

The easiest way to determine the transfer characteristic is to work backward; that is, for a given v_O we determine the corresponding value of v_I. We shall outline the process for positive v_O:

1. Assume a value for v_O.
2. Determine the load current i_L,

$$i_L = v_O/R_L$$

3. Use Eq. (12.27) to determine the current conducted by Q_N, i_N.
4. Determine v_{BEN} from

$$v_{BEN} = V_T \ln\left(i_N/I_S\right)$$

5. Determine v_I from

$$v_I = v_O + v_{BEN} - V_{BB}/2$$

It is also useful to find i_P and v_{EBP} as follows:

$$i_P = i_N - i_L$$
$$v_{EBP} = V_T \ \ln\left(i_P/I_S\right)$$

A similar process can be employed for negative v_O. However, symmetry can be utilized, obviating the need to repeat the calculations. The results obtained are displayed in the following table:

v_O (V)	i_L (mA)	i_N (mA)	i_P (mA)	v_{BEN} (V)	v_{EBP} (V)	v_I (V)	v_O/v_I	R_{out} (Ω)	v_o/v_i
+10.0	100	100.04	0.04	0.691	0.495	10.1	0.99	0.25	1.00
+5.0	50	50.08	0.08	0.673	0.513	5.08	0.98	0.50	1.00
+1.0	10	10.39	0.39	0.634	0.552	1.041	0.96	2.32	0.98
+0.5	5	5.70	0.70	0.619	0.567	0.526	0.95	4.03	0.96
+0.2	2	3.24	1.24	0.605	0.581	0.212	0.94	5.58	0.95
+0.1	1	2.56	1.56	0.599	0.587	0.106	0.94	6.07	0.94
0	0	2	2	0.593	0.593	0	—	6.25	0.94
–0.1	–1	1.56	2.56	0.587	0.599	–0.106	0.94	6.07	0.94
–0.2	–2	1.24	3.24	0.581	0.605	–0.212	0.94	5.58	0.95
–0.5	–5	0.70	5.70	0.567	0.619	–0.526	0.95	4.03	0.96
–1.0	–10	0.39	10.39	0.552	0.634	–1.041	0.96	2.32	0.98
–5.0	–50	0.08	50.08	0.513	0.673	–5.08	0.98	0.50	1.00
–10.0	–100	0.04	100.04	0.495	0.691	–10.1	0.99	0.25	1.00

The table also provides values for the dc gain v_O/v_I as well as the incremental gain v_o/v_i at the various values of v_O. The incremental gain is computed as follows

$$\frac{v_o}{v_i} = \frac{R_L}{R_L + R_{out}}$$

where R_{out} is the small-signal output resistance of the amplifier, given by Eq. (12.31). The incremental gain is the slope of the voltage-transfer characteristic, and the magnitude of its variation over the range of v_O is an indication of the linearity of the output stage. Observe that for $0 \leq |v_O| \leq 10$ V, the incremental gain changes from 0.94 to 1.00, about 6%. Also observe as v_O becomes positive, Q_N supplies more and more of i_L and the current in Q_P is correspondingly reduced. The opposite happens for negative v_O.

EXERCISE

12.5 To increase the linearity of the class AB output stage, the quiescent current I_Q is increased. The price paid is an increase in quiescent power dissipation. For the output stage considered in Example 12.3:

(a) Find the quiescent power dissipation.

(b) If I_Q is increased to 10 mA, find v_o/v_i at $v_O = 0$ and at $|v_O| = 10$ V, and hence the percentage change. Compare to the case in Example 12.3.

(c) Find the quiescent power dissipation for the case in (b).

Ans. (a) 60 mW; (b) 0.988 to 1.00; for a change of 1.2% compared to the 6% change in Example 12.3; (c) 300 mW

12.5 Biasing the Class AB Circuit

In this section we discuss two approaches for generating the voltage V_{BB} required for biasing the class AB output stage.

12.5.1 Biasing Using Diodes

Figure 12.14 shows a class AB circuit in which the bias voltage V_{BB} is generated by passing a constant current I_{BIAS} through a pair of diodes, or diode-connected transistors, D_1 and D_2. In circuits that supply large amounts of power, the output transistors are large-geometry devices. The biasing diodes, however, need not be large devices, and thus the quiescent current I_Q established in Q_N and Q_P will be $I_Q = nI_{\text{BIAS}}$, where n is the ratio of the emitter–junction area of the output devices to the junction area of the biasing diodes. In other words, the saturation (or scale) current I_S of the output transistors is n times that of the biasing diodes. Area ratioing is simple to implement in integrated circuits but difficult to realize in discrete-circuit designs.

When the output stage of Fig. 12.14 is sourcing current to the load, the base current of Q_N increases from I_Q/β_N (which is usually small) to approximately i_L/β_N. This base current drive must be supplied by the current source I_{BIAS}. It follows that I_{BIAS} must be greater than the maximum anticipated base drive for Q_N. This sets a lower limit on the value of I_{BIAS}. Now, since $n = I_Q/I_{BIAS}$, and since I_Q is usually much smaller than the peak load current ($<10\%$), we see that we cannot make n a large number. In other words, we cannot make the diodes much smaller than the output devices. This is a disadvantage of the diode biasing scheme.

From the discussion above we see that the current through the biasing diodes will decrease when the output stage is sourcing current to the load. Thus the bias voltage V_{BB} will also decrease, and the analysis of Section 12.4 must be modified to take this effect into account.

The diode biasing arrangement has an important advantage: It can provide thermal stabilization of the quiescent current in the output stage. To appreciate this point, recall that the class AB output stage dissipates power under quiescent conditions. Power dissipation

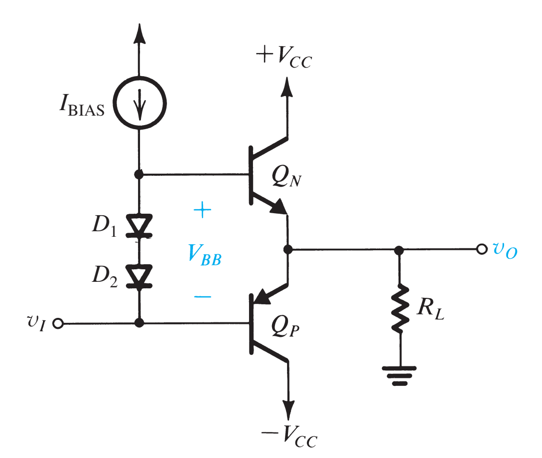


Figure 12.14 A class AB output stage utilizing diodes for biasing. If the junction area of the output devices, Q_N and Q_P, is n times that of the biasing devices D_1 and D_2, a quiescent current $I_Q = nI_{\text{BIAS}}$ flows in the output devices.

raises the internal temperature of the BJTs. From Chapter 6 we know that a rise in transistor temperature results in a decrease in its V_{BE} (approximately –2 mV/°C) if the collector current is held constant. Alternatively, if V_{BE} is held constant and the temperature increases, the collector current increases. The increase in collector current increases the power dissipation, which in turn increases the junction temperature and hence, once more, the collector current. Thus a positive-feedback mechanism exists that can result in a phenomenon called **thermal runaway**. Unless checked, thermal runaway can lead to the ultimate destruction of the BJT. Diode biasing can be arranged to provide a compensating effect that can protect the output transistors against thermal runaway under quiescent conditions. Specifically, if the diodes are in close thermal contact with the output transistors, their temperature will increase by the same amount as that of Q_N and Q_P. Thus V_{BB} will decrease at the same rate as $V_{BEN} + V_{EBP}$, with the result that I_Q remains constant. Close thermal contact is easily achieved in IC fabrication. It is obtained in discrete circuits by mounting the bias diodes on the metal case of Q_N or Q_P. Finally, it is important to note that thermal runaway does not occur in MOS circuits.

Example 12.4

Consider the class AB output stage under the conditions that $V_{CC} = 15$ V, $R_L = 100\ \Omega$, and the output is sinusoidal with a maximum amplitude of 10 V. Let Q_N and Q_P be matched with $I_S = 10^{-13}$ A and $\beta = 50$. Assume that the biasing diodes have one-third the junction area of the output devices. Find the value of I_{BIAS} that guarantees a minimum of 1 mA through the diodes at all times. Determine the quiescent current and the quiescent power dissipation in the output transistors (i.e., at $v_O = 0$). Also find V_{BB} for $v_O = 0$, +10 V, and −10 V.

Solution

The maximum current through Q_N is approximately equal to $i_{L\max} = 10\text{ V}/0.1\text{ k}\Omega = 100$ mA. Thus the maximum base current in Q_N is approximately 2 mA. To maintain a minimum of 1 mA through the diodes, we select $I_{\text{BIAS}} = 3$ mA. The area ratio of 3 yields a quiescent current of 9 mA through Q_N and Q_P. The quiescent power dissipation is

$$P_{DQ} = 2 \times 15 \times 9 = 270\text{ mW}$$

For $v_O = 0$, the base current of Q_N is $9/51 \simeq 0.18$ mA, leaving a current of $3 - 0.18 = 2.82$ mA to flow through the diodes. Since the diodes have $I_S = \frac{1}{3} \times 10^{-13}$ A, the voltage V_{BB} will be

$$V_{BB} = 2V_T \ln\left(\frac{2.82\text{ mA}}{I_S}\right) = 1.26\text{ V}$$

At $v_O = +10$ V, the current through the diodes will decrease to 1 mA, resulting in $V_{BB} \simeq 1.21$ V. At the other extreme of $v_O = -10$ V, Q_N will be conducting a very small current; thus its base current will be negligibly small and all of I_{BIAS} (3 mA) flows through the diodes, resulting in $V_{BB} \simeq 1.26$ V.

EXERCISES

12.6 For the circuit of Example 12.4, find i_N and i_P for $v_O = +10$ V and $v_O = -10$ V. (*Hint*: Use the V_{BB} values found in Example 12.4.)
Ans. 100.1 mA, 0.1 mA; 0.8 mA, 100.8 mA

12.7 If the collector current of a transistor is held constant, its v_{BE} decreases by 2 mV for every 1°C rise in temperature. Alternatively, if v_{BE} is held constant, then i_C increases by approximately $g_m \times 2$ mV for every 1°C rise in temperature. For a device operating at $I_C = 10$ mA, find the change in collector current resulting from an increase in temperature of 5°C.
Ans. 4 mA

12.5.2 Biasing Using the V_{BE} Multiplier

An alternative biasing arrangement that provides the designer with considerably more flexibility in both discrete and integrated designs is shown in Fig. 12.15. The bias circuit consists of transistor Q_1 with a resistor R_1 connected between base and emitter and a feedback resistor R_2 connected between collector and base. The resulting two-terminal network is fed with a constant-current source I_{BIAS}. If we neglect the base current of Q_1, then R_1 and R_2 will carry the same current I_R, given by

$$I_R = \frac{V_{BE1}}{R_1} \tag{12.32}$$

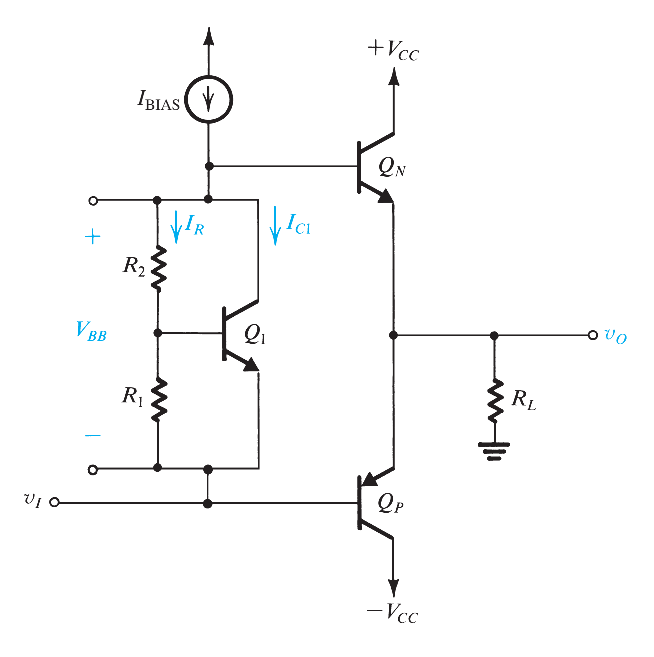


Figure 12.15 A class AB output stage utilizing a V_{BE} multiplier for biasing.

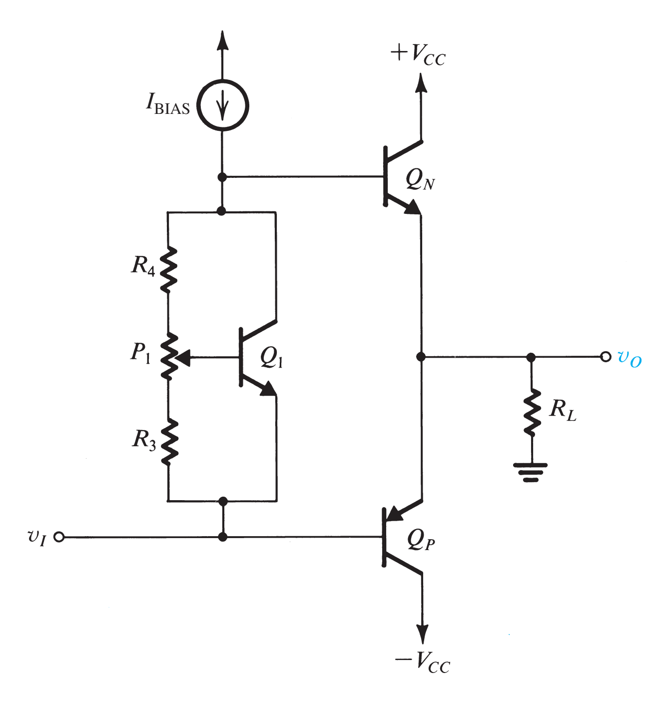


Figure 12.16 A discrete-circuit class AB output stage with a potentiometer used in the V_{BE} multiplier. The potentiometer is adjusted to yield the desired value of quiescent current in Q_N and Q_P.

and the voltage V_{BB} across the bias network will be

$$V_{BB} = I_R(R_1 + R_2)$$

$$= V_{BE1}\left(1 + \frac{R_2}{R_1}\right) \tag{12.33}$$

Thus the circuit simply multiplies V_{BE1} by the factor $(1 + R_2/R_1)$ and is known as the "V_{BE} multiplier." The multiplication factor is obviously under the designer's control and can be used to establish the value of V_{BB} required to yield a desired quiescent current I_Q. In IC design it is relatively easy to control accurately the ratio of two resistances. In discrete-circuit design, a potentiometer can be used, as shown in Fig. 12.16, and is manually set to produce the desired value of I_Q.

The value of V_{BE1} in Eq. (12.33) is determined by the portion of I_{BIAS} that flows through the collector of Q_1; that is,

$$I_{C1} = I_{BIAS} - I_R \tag{12.34}$$

$$V_{BE1} = V_T \ln \frac{I_{C1}}{I_{S1}} \tag{12.35}$$

where we have neglected the base current of Q_N, which is normally small both under quiescent conditions and when the output voltage is swinging negative. However, for positive v_O, especially at and near its peak value, the base current of Q_N can become sizable and will

reduce the current available for the V_{BE} multiplier. Nevertheless, since large changes in I_{C1} correspond to only small changes in V_{BE1}, the decrease in current will be mostly absorbed by Q_1, leaving I_R, and hence V_{BB}, almost constant.

EXERCISE

12.8 Consider a V_{BE} multiplier with $R_1 = R_2 = 1.2\ \text{k}\Omega$, utilizing a transistor that has $V_{BE} = 0.6$ V at $I_C = 1$ mA, and a very high β. (a) Find the value of the current I that should be supplied to the multiplier to obtain a terminal voltage of 1.2 V. (b) Find the value of I that will result in the terminal voltage changing (from the 1.2-V value) by +50 mV, +100 mV, +200 mV, –50 mV, –100 mV, –200 mV.
Ans. (a) 1.5 mA; (b) 3.24 mA, 7.93 mA, 55.18 mA, 0.85 mA, 0.59 mA, 0.43 mA

Like the diode biasing network, the V_{BE}–multiplier circuit can provide thermal stabilization of I_Q. This is especially true if $R_1 = R_2$, and Q_1 is in close thermal contact with the output transistors.

Example 12.5

It is required to redesign the output stage of Example 12.4 utilizing a V_{BE} multiplier for biasing. Use a small-geometry transistor for Q_1 with $I_S = 10^{-14}$ A and design for a quiescent current $I_Q = 2$ mA.

Solution

Since the peak positive current is 100 mA, the base current of Q_N can be as high as 2 mA. We shall therefore select $I_{\text{BIAS}} = 3$ mA, thus providing the multiplier with a minimum current of 1 mA.

Under quiescent conditions ($v_O = 0$ and $i_L = 0$) the base current of Q_N can be neglected and all of I_{BIAS} flows through the multiplier. We now must decide on how this current (3 mA) is to be divided between I_{C1} and I_R. If we select I_R greater than 1 mA, the transistor will be almost cut off at the positive peak of v_O. Therefore, we shall select $I_R = 0.5$ mA, leaving 2.5 mA for I_{C1}.

To obtain a quiescent current of 2 mA in the output transistors, V_{BB} should be

$$V_{BB} = 2V_T \ln \frac{2 \times 10^{-3}}{10^{-13}} = 1.19 \text{ V}$$

We can now determine $R_1 + R_2$ as follows:

$$R_1 + R_2 = \frac{V_{BB}}{I_R} = \frac{1.19}{0.5} = 2.38 \text{ k}\Omega$$

At a collector current of 2.5 mA, Q_1 has

$$V_{BE1} = V_T \ln \frac{2.5 \times 10^{-3}}{10^{-14}} = 0.66 \text{ V}$$

The value of R_1 can now be determined as

$$R_1 = \frac{0.66}{0.5} = 1.32 \text{ k}\Omega$$

and R_2 as

$$R_2 = 2.38 - 1.32 = 1.06 \text{ k}\Omega$$

12.6 Variations on the Class AB Configuration



In this section, we discuss a number of circuit improvements and protection techniques for the BJT class AB output stage.

12.6.1 Use of Input Emitter Followers

Figure 12.17 shows a class AB circuit biased using transistors Q_1 and Q_2, which also function as emitter followers, thus providing the circuit with a high input resistance. In effect, the

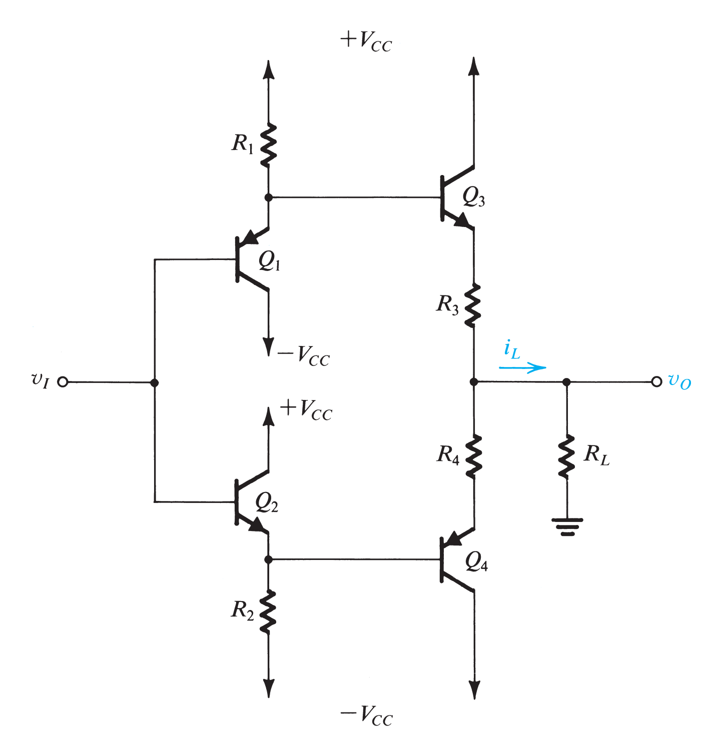


Figure 12.17 A class AB output stage with an input buffer. In addition to providing a high input resistance, the buffer transistors Q_1 and Q_2 bias the output transistors Q_3 and Q_4.

Q_1–Q_2 circuit functions as a unity-gain buffer amplifier. Since all four transistors are usually matched, and neglecting the effect of R_3 and R_4, we see that the quiescent current ($v_I = 0$, $R_L = \infty$) in Q_3 and Q_4 is equal to that in Q_1 and Q_2. Resistors R_3 and R_4 are usually very small and are included to compensate for possible mismatches between Q_3 and Q_4 and to guard against the possibility of thermal runaway due to temperature differences between the input- and output-stage transistors. The latter point can be appreciated by noting that an increase in the current of, say, Q_3 causes an increase in the voltage drop across R_3 and a corresponding decrease in V_{BE3}. Thus R_3 provides negative feedback that helps stabilize the current through Q_3.

Because the circuit of Fig. 12.17 requires high-quality *pnp* transistors, it is not suitable for implementation in conventional monolithic IC technology. However, excellent results have been obtained with this circuit implemented in hybrid thick-film technology (Wong and Sherwin, 1979). This technology permits component trimming, for instance, to minimize the output offset voltage. The circuit can be used alone or together with an op amp to provide increased output driving capability.

EXERCISE

12.9 (*Note:* Although rather long, this exercise is very instructive.) Consider the circuit of Fig. 12.17 with $R_1 = R_2 = 5\ \text{k}\Omega$, $R_3 = R_4 = 0\ \Omega$, and $V_{CC} = 15$ V. Let the transistors be matched with $I_S = 3.3 \times 10^{-14}$ A and $\beta = 200$. (These are the values used in the LH002 manufactured by National Semiconductor, except that $R_3 = R_4 = 2\ \Omega$ there.) (a) For $v_I = 0$ and $R_L = \infty$, find the quiescent current in each of the four transistors and v_O. (b) For $R_L = \infty$, find i_{C1}, i_{C2}, i_{C3}, i_{C4}, and v_O for $v_I = +10$ V and -10 V. (c) Repeat (b) for $R_L = 100\ \Omega$.

Ans. (a) 2.87 mA; 0 V; (b) for $v_I = +10$ V: 0.88 mA, 4.87 mA, 1.95 mA, 1.95 mA, +9.98 V; for $v_I = -10$ V: 4.87 mA, 0.88 mA, 1.95 mA, 1.95 mA, −9.98 V; (c) for $v_I = +10$ V: 0.38 mA, 4.87 mA, 100 mA, 0.02 mA, +9.86 V; for $v_I = -10$ V: 4.87 mA, 0.38 mA, 0.02 mA, 100 mA, −9.86 V

12.6.2 Use of Compound Devices

To increase the current gain of the output-stage transistors, and thus reduce the required base current drive, the Darlington configuration shown in Fig. 12.18 is frequently used to replace the *npn* transistor of the class AB stage. The Darlington configuration is equivalent to a single *npn* transistor having $\beta \simeq \beta_1\beta_2$, but almost twice the value of V_{BE}.

The Darlington configuration can be also used for *pnp* transistors, and this is indeed done in discrete-circuit design. In IC design, however, the lack of good-quality *pnp* transistors prompted the use of the alternative compound configuration shown in Fig. 12.19. This compound device is equivalent to a single *pnp* transistor having $\beta \simeq \beta_1\beta_2$. When fabricated with standard IC technology, Q_1 is usually a lateral *pnp* having a low β ($\beta = 5 - 10$) and poor high-frequency response ($f_T \simeq 5$ MHz); see Appendix A and Appendix K. The compound

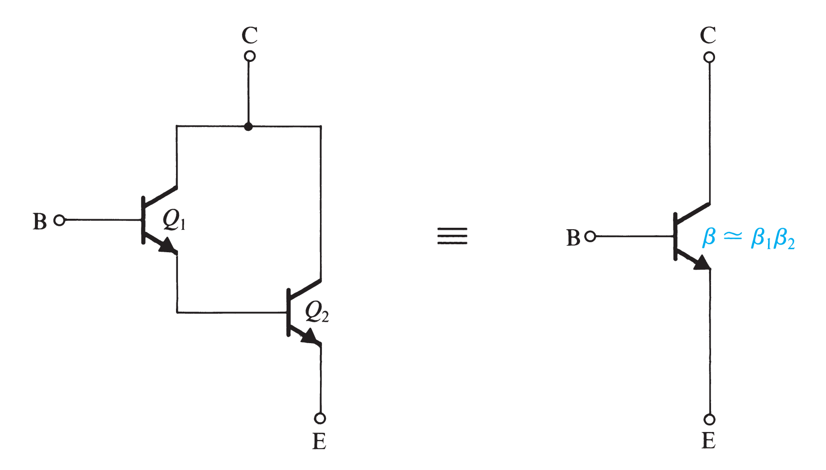


Figure 12.18 The Darlington configuration.

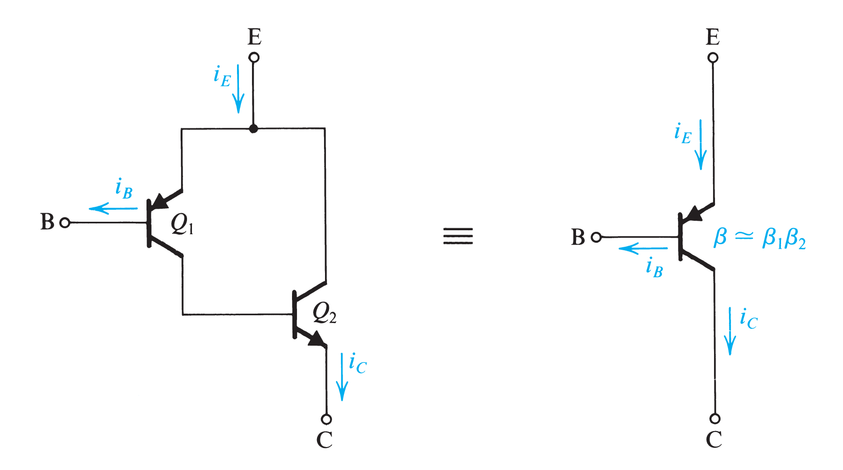


Figure 12.19 The compound-*pnp* configuration.

device, although it has a relatively high equivalent β, still suffers from a poor high-frequency response. It also suffers from another problem: The feedback loop formed by Q_1 and Q_2 is prone to high-frequency oscillations (with frequency near f_T of the *pnp* device, i.e., about 5 MHz). Methods exist for preventing such oscillations. The subject of feedback-amplifier stability was studied in Chapter 11.

To illustrate the application of the Darlington configuration and of the compound *pnp,* we show in Fig. 12.20 an output stage utilizing both. Class AB biasing is achieved using a V_{BE} multiplier. Note that the Darlington *npn* adds one more V_{BE} drop, and thus the V_{BE} multiplier is required to provide a bias voltage of about 2 V. The design of this class AB stage is investigated in Problem 12.39.

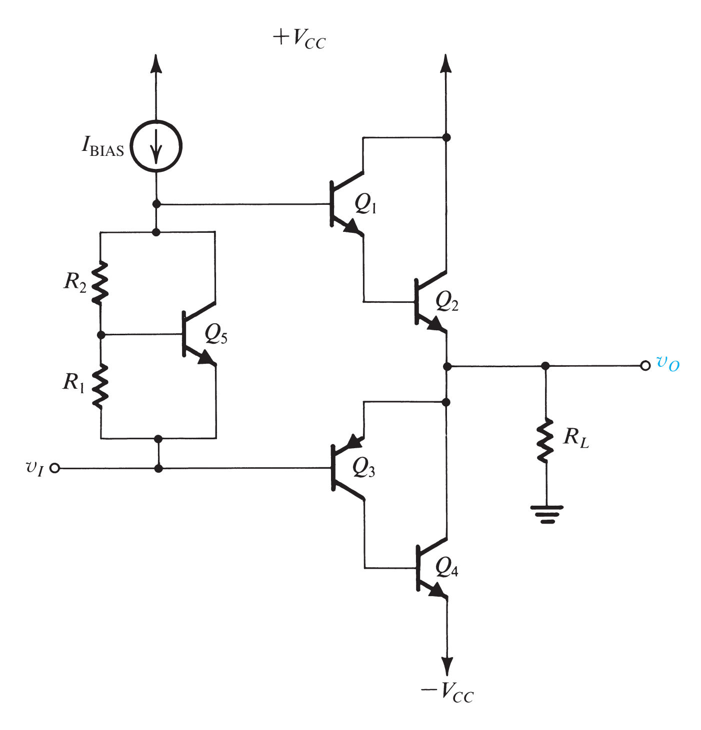


Figure 12.20 A class AB output stage utilizing a Darlington *npn* and a compound *pnp*. Biasing is obtained using a V_{BE} multiplier.

EXERCISE

12.10 (a) Refer to Fig. 12.19. Show that, for the composite *pnp* transistor,

$$i_B \simeq \frac{i_C}{\beta_N \beta_P}$$

and

$$i_E \simeq i_C$$

Hence show that

$$i_C \simeq \beta_N I_{SP} e^{v_{EB}/V_T}$$

and thus the transistor has an effective scale current

$$I_S = \beta_N I_{SP}$$

where I_{SP} is the scale current of the *pnp* transistor Q_1.
(b) For $\beta_P = 20$, $\beta_N = 50$, $I_{SP} = 10^{-14}$ A, find the effective current gain of the compound device and its v_{EB} when $i_C = 100$ mA.
Ans. (b) 1000; 0.651 V

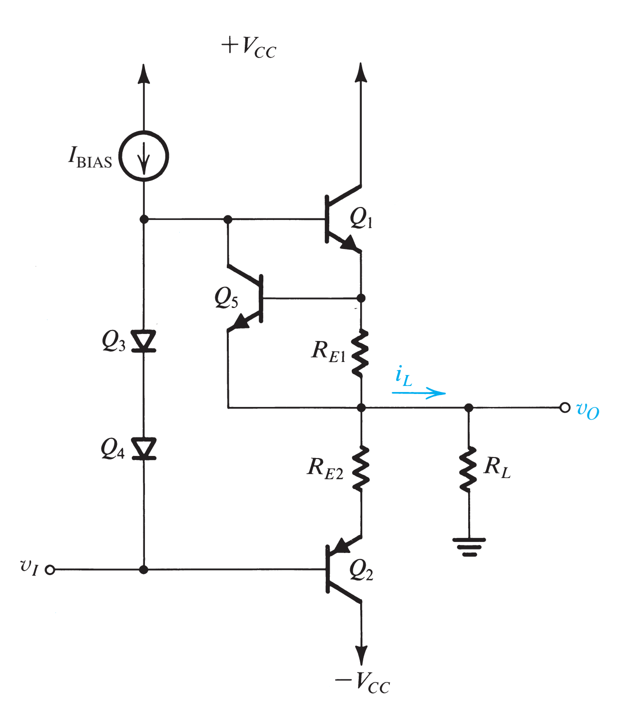


Figure 12.21 A class AB output stage with short-circuit protection. The protection circuit shown operates in the event of an output short circuit while v_O is positive.

12.6.3 Short-Circuit Protection

Figure 12.21 shows a class AB output stage equipped with protection against the effect of short-circuiting the output while the stage is sourcing current. The large current that flows through Q_1 in the event of a short circuit will develop a voltage drop across R_{E1} of sufficient value to turn Q_5 on. The collector of Q_5 will then conduct most of the current I_{BIAS}, robbing Q_1 of its base drive. The current through Q_1 will thus be reduced to a safe operating level.

This method of short-circuit protection is effective in ensuring device safety, but it has the disadvantage that under normal operation about 0.5 V drop might appear across each R_E. This means that the voltage swing at the output will be reduced by that much, in each direction. On the other hand, the inclusion of emitter resistors provides the additional benefit of protecting the output transistors against thermal runaway.

EXERCISE

D12.11 In the circuit of Fig. 12.21 let $I_{BIAS} = 2$ mA. Find the value of R_{E1} that causes Q_5 to turn on and absorb all 2 mA when the output current being sourced reaches 150 mA. For Q_5, $I_S = 10^{-14}$ A. If the normal peak output current is 100 mA, find the voltage drop across R_{E1} and the collector current of Q_5.
Ans. 4.3 Ω; 430 mV; 0.3 μA

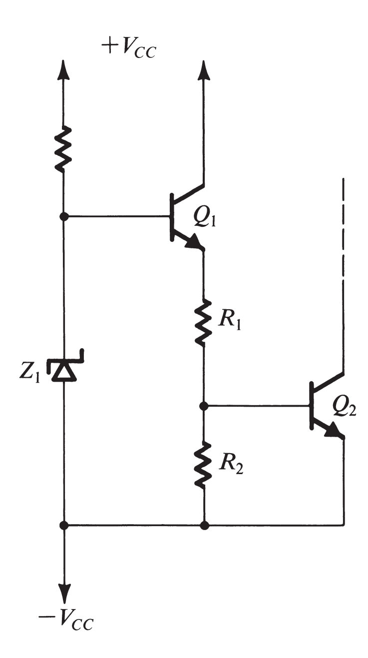


Figure 12.22 Thermal-shutdown circuit.

12.6.4 Thermal Shutdown

In addition to short-circuit protection, most IC power amplifiers are usually equipped with a circuit that senses the temperature of the chip and turns on a transistor in the event that the temperature exceeds a safe preset value. The turned-on transistor is connected in such a way that it absorbs the bias current of the amplifier, thus virtually shutting down its operation.

Figure 12.22 shows a thermal-shutdown circuit. Here, transistor Q_2 is normally off. As the chip temperature rises, the combination of the positive temperature coefficient of zener diode Z_1 and the negative temperature coefficient of V_{BE1} causes the voltage at the emitter of Q_1 to rise. This in turn raises the voltage at the base of Q_2 to the point at which Q_2 turns on.

12.7 CMOS Class AB Output Stages

In this section we study CMOS class AB output stages. We begin with the CMOS counterpart of the BJT class AB output stage studied in Section 12.5. As we shall see, this circuit suffers from a relatively low output signal swing, a serious limitation, especially in view of the shrinking power-supply voltages characteristic of modern deep-submicron CMOS technologies. We will then look at an attractive alternative circuit that overcomes this problem.

12.7.1 The Classical Configuration

Figure 12.23 shows the classical CMOS class AB output stage. The circuit is the exact counterpart of the bipolar circuit shown in Fig. 12.14 with the biasing diodes implemented with diode-connected transistors Q_1 and Q_2. The constant current I_{BIAS} flowing through Q_1 and Q_2 establishes a dc bias voltage V_{GG} between the gates of Q_N and Q_P. This voltage in turn establishes the quiescent ($v_O = 0$) current I_Q in Q_N and Q_P. Unlike the BJT circuit in Fig. 12.14, here the zero dc gate current of Q_N results in the current through Q_1 and Q_2 remaining constant

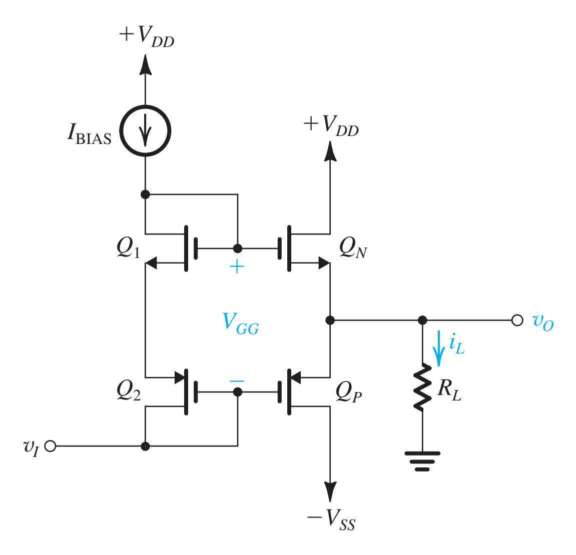


Figure 12.23 Classical CMOS class AB output stage. This circuit is the CMOS counterpart of the BJT circuit in Fig. 11.14 with the biasing diodes implemented with diode-connected MOSFETs Q_1 and Q_2

at I_{BIAS} irrespective of the value of v_O and the load current i_L. Thus V_{GG} remains constant and the circuit is more like the idealized bipolar case shown in Fig. 12.11.

The value of I_Q can be determined by utilizing the i_D–v_{GS} equations for the four MOS transistors for the case $v_O = 0$. Neglecting channel-length modulation, we can write for Q_1,

$$I_{D1} = I_{\text{BIAS}} = \frac{1}{2}k_n'(W/L)_1(V_{GS1} - V_{tn})^2 \tag{12.36}$$

and for Q_2,

$$I_{D2} = I_{\text{BIAS}} = \frac{1}{2}k_p'(W/L)_2(V_{SG2} - |V_{tp}|)^2 \tag{12.37}$$

Equations (12.36) and (12.37) can be used to find V_{GS1} and V_{SG2}, which when summed yield V_{GG}; thus,

$$V_{GG} = V_{GS1} + V_{SG2} = V_{tn} + |V_{tp}| + \sqrt{2I_{\text{BIAS}}}\left(\frac{1}{\sqrt{k_n'(W/L)_1}} + \frac{1}{\sqrt{k_p'(W/L)_2}}\right) \tag{12.38}$$

We can follow a similar process for Q_N and Q_P, which, for $v_O = 0$, are conducting the quiescent current I_Q; thus,

$$V_{GG} = V_{GSN} + V_{SGP} = V_{tn} + |V_{tp}| + \sqrt{2I_Q}\left(\frac{1}{\sqrt{k_n'(W/L)_n}} + \frac{1}{\sqrt{k_p'(W/L)_p}}\right) \tag{12.39}$$

Equations (12.38) and (12.39) can be combined to obtain

$$I_Q = I_{\text{BIAS}}\left[\frac{1/\sqrt{k_n'(W/L)_1} + 1/\sqrt{k_p'(W/L)_2}}{1/\sqrt{k_n'(W/L)_n} + 1/\sqrt{k_p'(W/L)_p}}\right]^2 \tag{12.40}$$

which indicates that I_Q is determined by I_{BIAS} together with the (W/L) ratios of the four transistors. For the case Q_1 and Q_2 are matched, that is,

$$k_p'(W/L)_2 = k_n'(W/L)_1 \tag{12.41}$$

and Q_N and Q_P are matched, that is,

$$k_p'(W/L)_p = k_n'(W/L)_n \tag{12.42}$$

Equation (12.40) simplifies to

$$I_Q = I_{\text{BIAS}} \frac{(W/L)_n}{(W/L)_1} \tag{12.43}$$

which is an intuitively appealing result.

EXERCISE

12.12 For the CMOS class AB output stage of Fig. 12.23, consider the case of matched Q_1 and Q_2, and matched Q_N and Q_P. If $I_Q = 1$ mA and $I_{\text{BIAS}} = 0.2$ mA, find (W/L) for each of Q_1, Q_2, Q_N, and Q_P so that in the quiescent state each transistor operates at an overdrive voltage of 0.2 V. Let $V_{DD} = V_{SS} = 2.5$ V, $k_n' = 250\ \mu\text{A/V}^2$, $k_p' = 100\ \mu\text{A/V}^2$, and $V_{tn} = -V_{tp} = 0.5$ V. Also find V_{GG}.
Ans. 40; 100; 200; 500; 1.4 V

A drawback of the CMOS class AB circuit of Fig. 12.23 is the restricted range of output voltage swing. To find the maximum possible value of v_O, refer to Fig. 12.23 and assume that across the bias current source is a dc voltage of V_{BIAS}. We can write for v_O,

$$v_O = V_{DD} - V_{\text{BIAS}} - v_{GSN} \tag{12.44}$$

The maximum value of v_O will be limited by the need to keep V_{BIAS} to a minimum of V_{OV} of the transistor supplying I_{BIAS} (otherwise the current-source transistor no longer operates in saturation); thus,

$$v_{O\max} = V_{DD} - V_{OV}|_{\text{BIAS}} - v_{GSN} \tag{12.45}$$

Note that when v_O is at its maximum value, Q_N will be supplying most or all of i_L, and v_{GSN} will be large, thus

$$v_{O\max} = V_{DD} - V_{OV}|_{\text{BIAS}} - V_{tn} - v_{OVN} \tag{12.46}$$

where v_{OVN} is the overdrive voltage of Q_N when it is supplying $i_{L\max}$.

EXERCISE

12.13 For the circuit specified in Exercise 12.12, find $v_{O\max}$ when $i_{L\max} = 10$ mA. Assume that Q_N is supplying all of $i_{L\max}$ and that $V_{OV}|_{\text{BIAS}} = 0.2$ V.
Ans. 1.17 V

The minimum allowed value of v_O can be found in a similar way. Here we note that the transistor supplying v_I (not shown) will need a minimum voltage across it of $V_{OV}|_I$. Thus,

$$v_{O\min} = -V_{SS} + V_{OV}|_I + |V_{tp}| + |v_{OVP}| \tag{12.47}$$

where $|v_{OVP}|$ is the overdrive voltage of Q_P when sinking the maximum negative value of i_L.

Finally, we observe that the reason for the lower allowable range of v_O in the CMOS circuit is the relatively large value of v_{OVN} and $|v_{OVP}|$; that is, the large values of v_{GSN} and v_{SGP} required to supply the large output currents. In the BJT circuit the corresponding voltages, v_{BEN} and v_{EBP}, remain close to 0.7 V. The overdrive voltages v_{OVN} and $|v_{OVP}|$ can be reduced by making the W/L ratios of Q_N and Q_P large. This, however, can lead to impractically large devices.

12.7.2 An Alternative Circuit Utilizing Common-Source Transistors

The allowable range of v_O can be increased by replacing the source followers with a pair of complementary transistors connected in the common-source configuration, as shown in Fig. 12.24. Here Q_P supplies the load current when v_O is positive and allows v_O to go as high as $(V_{DD} - |v_{OVP}|)$, a much higher value than that given by Eq. (12.46). For negative v_O, Q_N sinks the load current and allows v_O to go as low as $-V_{SS} + v_{OVN}$. This also is larger in magnitude than the value given by Eq. (12.47). Thus, the circuit of Fig. 12.24 provides an output voltage

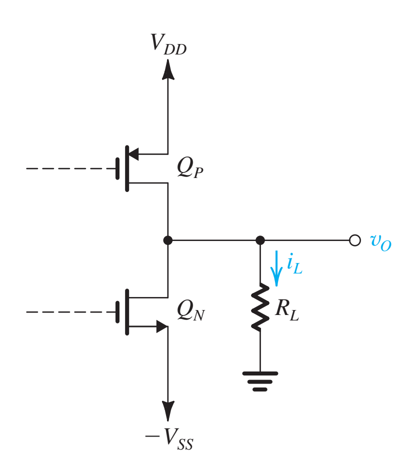


Figure 12.24 An alternative CMOS output stage utilizing a pair of complementary MOSFETs connected in the common-source configuration. The driving circuit is not shown.

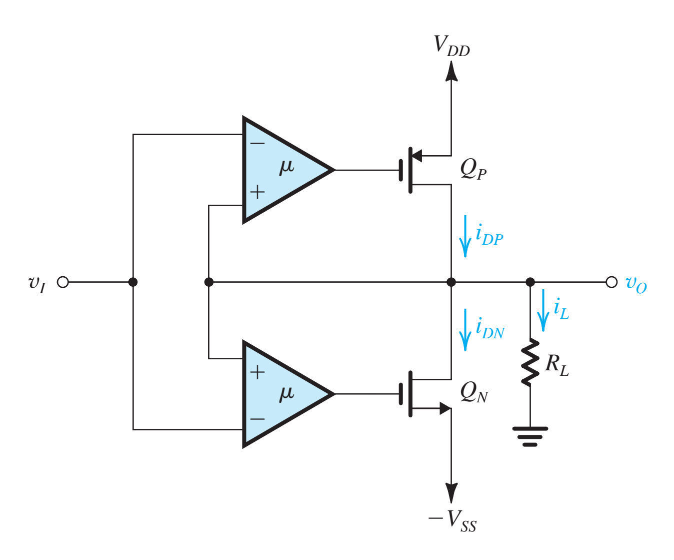


Figure 12.25 Inserting an amplifier in the negative feedback path of each of Q_N and Q_P reduces the output resistance and makes $v_O \simeq v_I$; both are desirable properties for the output stage.

range that is within an overdrive voltage of each of the supplies. The disadvantage of the circuit, however, is its high output resistance,

$$R_{\text{out}} = r_{on} \| r_{op} \tag{12.48}$$

To reduce the output resistance, negative feedback is employed as shown in Fig. 12.25. Here an amplifier with gain μ is inserted between drain and gate of each of Q_N and Q_P. To verify that the feedback around each amplifier is negative, assume that v_O increases. The top amplifier will cause the gate voltage of Q_P to increase, thus its v_{SG} decreases and i_{DP} decreases. The decrease in i_{DP} causes v_O to decrease, which is opposite to the initially assumed change, thus verifying that the feedback is negative. A similar process can be used to verify that the feedback around the bottom amplifier also is negative.

From our study of feedback in Chapter 11, we observe that each of the two feedback loops is of the series–shunt type, which is the topology appropriate for a voltage amplifier. Thus, as we shall show shortly, the feedback will reduce the output resistance of the amplifier. Also, observe that if the loop gain is large, the voltage difference between the two input terminals of each feedback amplifier, *the error voltage*, will be small, resulting in $v_O \simeq v_I$. For this reason, the two amplifiers μ are known as **error amplifiers**.

Both the low output resistance and the near-unity dc gain are highly desirable properties for an output stage.

Output Resistance To derive an expression for the output resistance R_{out}, we consider each half of the circuit separately, find its output resistance, $R_{\text{out}p}$ for the top half and $R_{\text{out}n}$ for the bottom half, and then obtain the overall output resistance as the parallel equivalent of the two resistances,

$$R_{\text{out}} = R_{\text{out}n} \| R_{\text{out}p} \tag{12.49}$$

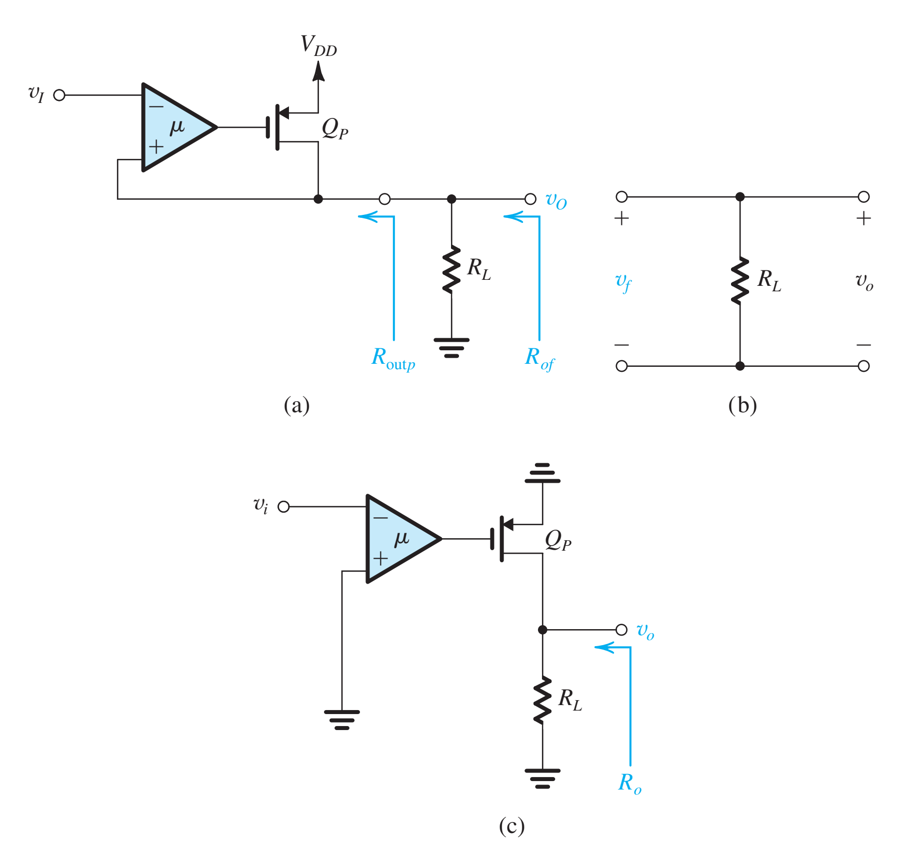


Figure 12.26 Determining the output resistance. **(a)** The top half of the output stage showing the definition of R_{outp} and R_{of}. **(b)** The β circuit and **(c)** the A circuit.

Figure 12.26(a) shows the top half of the circuit. Observe that feedback is applied by connecting the output back to the input. Thus the feedback network is the two-port shown in Fig. 12.26(b) and the feedback factor is

$$\beta = 1 \tag{12.50}$$

Including the loading effects of the feedback network results in the A circuit shown in Fig. 12.26(c). Note that since we are now interested in incremental quantities, we have replaced V_{DD} with a short circuit to ground. The open-loop gain A can be found from the circuit in Fig. 12.26(c) as

$$A \equiv \frac{v_o}{v_i} = \mu g_{mp}(r_{op} \| R_L) \tag{12.51}$$

where we have utilized implicitly the small-signal model of Q_P. The values of the small-signal parameters g_{mp} and r_{op} are to be evaluated at the current at which Q_P is operating. The open-loop output resistance R_o is found by inspection as

$$R_o = R_L \| r_{op} \tag{12.52}$$

The output resistance with feedback R_{of} can now be found as

$$R_{of} = \frac{R_o}{1+A\beta} = \frac{(R_L \| r_{op})}{1+\mu g_{mp}(r_{op} \| R_L)} \tag{12.53}$$

and the output resistance R_{outp} is found by excluding R_L from R_{of}, that is,

$$R_{outp} = 1\Big/\left(\frac{1}{R_{of}} - \frac{1}{R_L}\right) \tag{12.54}$$

which results in

$$R_{outp} = r_{op} \Big\| \frac{1}{\mu g_{mp}} \simeq \frac{1}{\mu g_{mp}} \tag{12.55}$$

which can be quite low. A similar development applied to the bottom half of the circuit in Fig. 12.25 results in

$$R_{outn} \simeq 1/\mu g_{mn} \tag{12.56}$$

Combining Eqs. (12.55) and (12.56) gives

$$R_{out} \simeq 1/\mu(g_{mp} + g_{mn}) \tag{12.57}$$

The Voltage-Transfer Characteristic Next we derive an expression for the voltage-transfer characteristic, v_O versus v_I, of the class AB common-source buffer. Toward that end, we first consider the circuit in the quiescent state, shown in Fig. 12.27(a). Here $v_I = 0$ and $v_O = 0$. Each of the error amplifiers is designed to deliver to the gate of its associated MOSFET the dc voltage required to establish the desired value of quiescent current I_Q.

Note that these voltages at the outputs of the error amplifiers are the dc bias or quiescent values. These voltages will undergo incremental changes when the input voltages of the amplifiers change from zero. To obtain class AB operation, I_Q is usually selected to be 10%

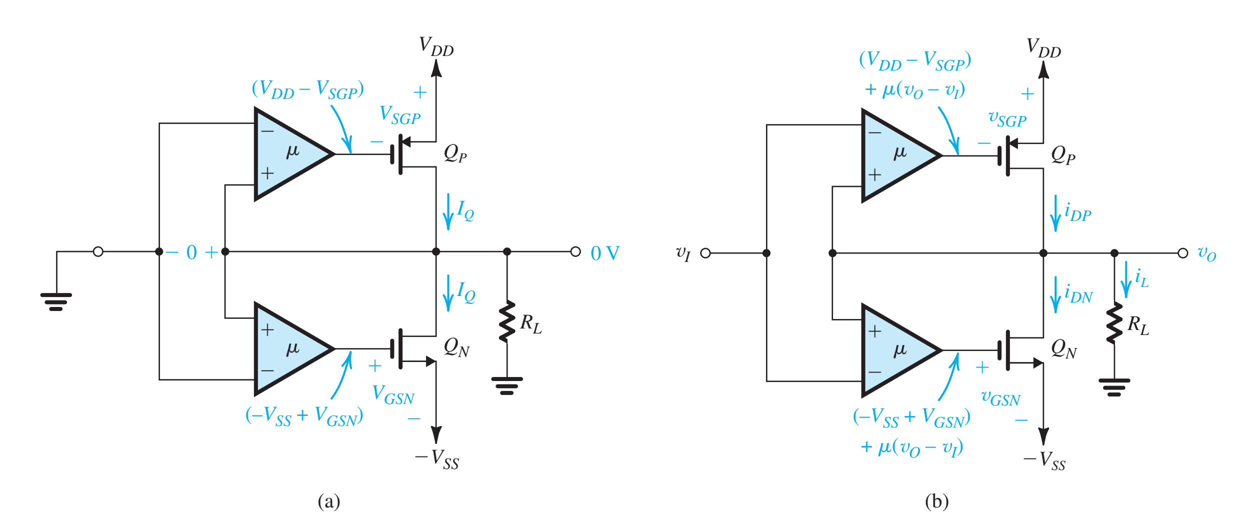


Figure 12.27 Analysis of the CMOS output stage to determine v_O versus v_I. **(a)** Quiescent conditions. **(b)** The situation with v_I applied.

or so of the maximum output current. Referring to Fig. 12.27(a), we can write for Q_P,

$$I_{DP} = I_Q = \frac{1}{2}k'_p\left(\frac{W}{L}\right)_p(V_{SGP} - |V_{tp}|)^2$$

Substituting $V_{SGP} = |V_{tp}| + V_{OV}$, where V_{OV} is the magnitude of the quiescent overdrive voltage of Q_P, gives

$$I_Q = \frac{1}{2}k'_p\left(\frac{W}{L}\right)_p V_{OV}^2 \tag{12.58}$$

Similarly, we obtain for Q_N

$$I_Q = \frac{1}{2}k'_n\left(\frac{W}{L}\right)_n V_{OV}^2 \tag{12.59}$$

Usually the two transistors are matched,

$$k'_p\left(\frac{W}{L}\right)_p = k'_n\left(\frac{W}{L}\right)_n = k$$

Thus,

$$I_Q = \frac{1}{2}kV_{OV}^2 \tag{12.60}$$

Next consider the situation with v_I applied, illustrated in Fig. 12.27(b). The voltage at the output of each of the error amplifiers increases by $\mu(v_O - v_I)$. Thus v_{SGP} decreases by $\mu(v_O - v_I)$ and v_{GSN} increases by $\mu(v_O - v_I)$, and we can write

$$\begin{aligned} i_{DP} &= \frac{1}{2}k[V_{OV} - \mu(v_O - v_I)]^2 \\ &= \frac{1}{2}kV_{OV}^2\left[1 - \mu\frac{v_O - v_I}{V_{OV}}\right]^2 \\ &= I_Q\left(1 - \mu\frac{v_O - v_I}{V_{OV}}\right)^2 \end{aligned} \tag{12.61}$$

and

$$i_{DN} = I_Q\left(1 + \mu\frac{v_O - v_I}{V_{OV}}\right)^2 \tag{12.62}$$

At the output node we have

$$i_L = i_{DP} - i_{DN} \tag{12.63}$$

Substituting for $i_L = v_O/R_L$ and for i_{DP} and i_{DN} from Eqs. (12.61) and (12.62), and solving the resulting equation to obtain v_O, results in

$$v_O = \frac{v_I}{1 + \dfrac{V_{OV}}{4\mu I_Q R_L}} \tag{12.64}$$

Usually $\left(V_{OV}/4\mu I_Q R_L\right) \ll 1$, enabling us to express v_O as

$$v_O \simeq v_I\left(1 - \frac{V_{OV}}{4\mu I_Q R_L}\right) \tag{12.65}$$

Thus the gain error is

$$\text{Gain error} \equiv \frac{v_O}{v_I} - 1 = -\frac{V_{OV}}{4\mu I_Q R_L} \tag{12.66}$$

Since at the quiescent point,

$$g_{mp} = g_{mn} = g_m = \frac{2I_Q}{V_{OV}} \tag{12.67}$$

the gain error can be expressed as

$$\text{Gain error} = -\frac{1}{2\mu g_m R_L} \tag{12.68}$$

Thus selecting a large value for μ results in reducing both the gain error and the output resistance. However, a large μ can make the quiescent current I_Q too dependent on the input offset voltages that are inevitably present in the error amplifiers. Typically, μ is selected in the range 5 to 10. Trade-offs are also present in the selection of I_Q: A large I_Q reduces crossover distortion, R_{out}, and gain error, at the expense of increased quiescent power dissipation.

Example 12.6

In this example we explore the design and operation of a class AB common-source output stage of the type shown in Fig. 12.25, required to operate from a ±2.5-V power supply to feed a load resistance $R_L = 100\,\Omega$. The transistors available have $V_{tn} = -V_{tp} = 0.5$ V and $k_n' = 2.5k_p' = 250\ \mu\text{A/V}^2$. The gain error is required to be less than 2.5% and $I_Q = 1$ mA.

Solution

The gain error is given by Eq. (12.66),

$$\text{Gain error} = -\frac{V_{OV}}{4\mu I_Q R_L}$$

We are given the required maximum gain error of -0.025, $I_Q = 1$ mA, and $R_L = 100\ \Omega$. In order to keep μ low and also obtain as high a g_m as possible $\left(g_m = 2I_Q/V_{OV}\right)$, we select V_{OV} to be as low as possible. Practically speaking, V_{OV} is usually 0.1 V to 0.2 V. Selecting $V_{OV} = 0.1$ V results in

$$0.025 = \frac{0.1}{4 \times \mu \times 1 \times 10^{-3} \times 100}$$

which yields

$$\mu = 10$$

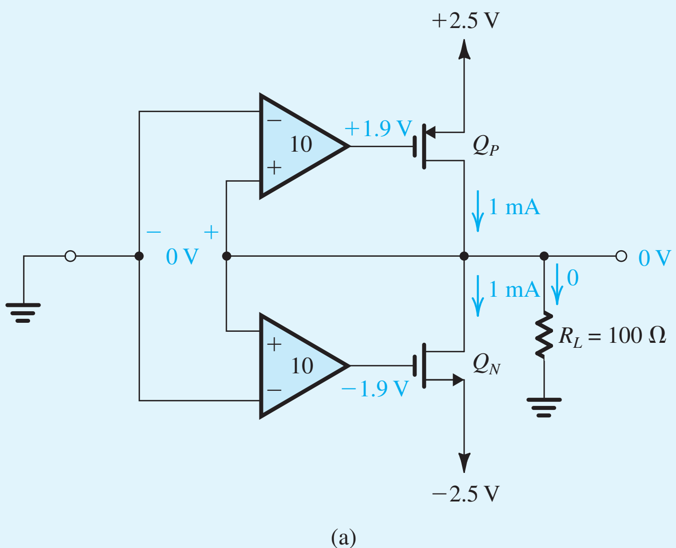
which is within the typically recommended range.

Figure 12.28(a) shows the circuit in the quiescent state with the various dc voltages and currents indicated. Note that the dc voltages (±1.9 V) at the output of the error amplifiers are the dc bias values,

obtained with zero input signals applied at the amplifier inputs. The required (W/L) ratios of Q_N and Q_P can be found as follows:

$$I_Q = \frac{1}{2} k'_p \left(\frac{W}{L}\right)_p V_{OV}^2$$

$$1 \times 10^{-3} = \frac{1}{2} \times 0.1 \times 10^{-3} \left(\frac{W}{L}\right)_p \times (0.1)^2$$



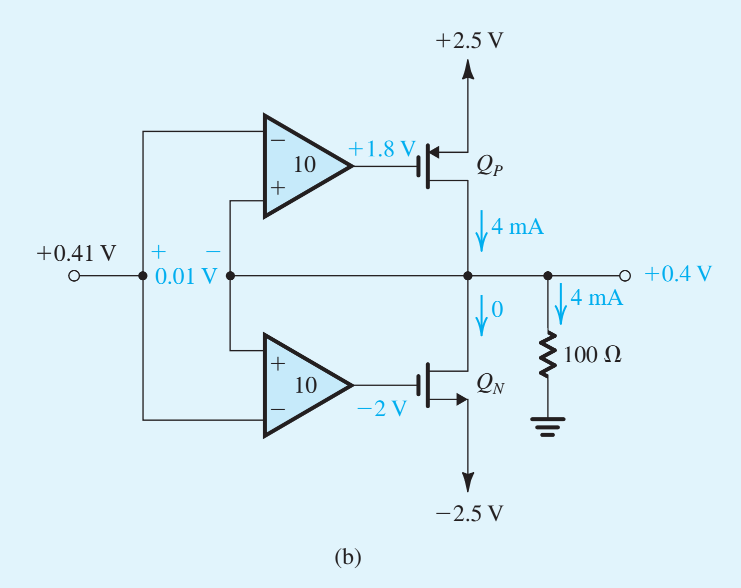


Figure 12.28 **(a)** Circuit in the quiescent state; **(b)** circuit at the point at which Q_N turns off; **(c)** conditions at $v_O = v_{O\max}$.

Example 12.6 *continued*

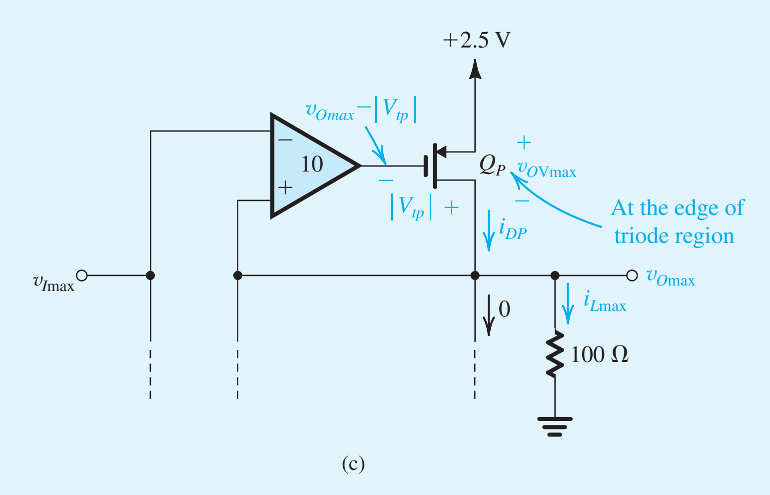


Figure 12.28 *continued*

Thus,

$$\left(\frac{W}{L}\right)_p = 2000$$

$$\left(\frac{W}{L}\right)_n = \frac{(W/L)_p}{k_n'/k_p'} = \frac{2000}{2.5} = 800$$

Thus Q_N and Q_P are very large transistors, not an unusual situation in a high-power output stage.

To obtain the output resistance at the quiescent point, we use Eq. (12.57),

$$R_{\text{out}} = \frac{1}{\mu(g_{mp} + g_{mn})}$$

where

$$g_{mp} = g_{mn} = \frac{2I_Q}{V_{OV}} = \frac{2 \times 1}{0.1} = 20\,\text{mA/V}$$

Thus

$$R_{\text{out}} = \frac{1}{10(0.02 + 0.02)} = 2.5\ \Omega$$

Next we wish to determine the maximum and minimum allowed values of v_O. Since the circuit is symmetrical, we need to consider only the positive-output or the negative-output case. For v_O positive, Q_P conducts more of the output current i_L. Eventually, Q_N turns off and Q_P conducts all of i_L. To find

the value of v_O at which this occurs, note that Q_N turns off when the voltage at its gate drops from the quiescent value of −1.9 V [see Fig. 12.28(a)] to −2 V, at which point $v_{GSN} = V_{tn}$. An equal change of −0.1 V appears at the output of the top amplifier, as shown in Fig. 12.28(b). Analysis of the circuit in Fig. 12.28(b) shows that

$$i_{DP} = \frac{1}{2} \times 0.1 \times 10^{-3} \times 2000 \times (0.2)^2 = 4 \text{ mA}$$

$$i_L = i_{DP} = 4 \text{ mA}$$

$$v_O = i_L R_L = 4 \times 10^{-3} \times 100 = 0.4 \text{ V}$$

$$v_I = v_O + (0.1/10) = 0.41 \text{ V}$$

For $v_O > 0.4$ V, Q_P must conduct all the current i_L. The situation at $v_O = v_{O\max}$ is illustrated in Fig. 12.28(c). Analysis of this circuit results in

$$v_{O\max} \simeq 2.05 \text{ V}$$

and

$$i_{L\max} = 20.5 \text{ mA}$$

EXERCISE

12.14 Suppose it is required to reduce the W/L ratios of Q_N and Q_P in the circuit considered in the above example by a factor of 2 while keeping I_Q at 1 mA. What value should be used for V_{OV}? What is the new value for the gain error and for R_{out} at the quiescent point?
Ans. 0.14 V; −3.5%; 3.5 Ω

12.8 IC Power Amplifiers



A variety of IC power amplifiers are available. Most consist of a high-gain, small-signal amplifier followed by a class AB output stage. Some have overall negative feedback already applied, resulting in a fixed closed-loop voltage gain. Others do not have on-chip feedback and are, in effect, op amps with large output-power capability. In fact, the output current-driving capability of any general-purpose op amp can be increased by cascading it with a class B or class AB output stage and applying overall negative feedback. The additional output stage can be either a discrete circuit or a hybrid IC such as the buffer discussed in Section 12.6. In the following we discuss some power-amplifier examples.

EARLY POWER-OP-AMP PRODUCT:

In 1985 Robert J. Widlar (1937–1991) and Mineo Yamatake at National Semiconductor introduced the LM12, probably the first very-high-power monolithic operational amplifier, offering an order-of-magnitude improvement over its predecessors. Nominally rated at 150-W output, this op amp could sustain 90 W of continuous sine-wave output with a 40-Ω load, while handling up to 800 W of short-term dynamic loading. The design operated from ±35-V supplies to provide a ±25-V signal with a ±10-A output. This monolithic amplifier employed polycrystalline film resistors for thermal stability. It incorporated a variety of novel protection features involving disconnection of the load from the output terminal, turn-on delay while awaiting internal stabilization, overtemperature control, and output-current limiting. It was internally unity-gain compensated with a unity-gain bandwidth of 700 kHz. While now obsolete, the LM12 was a clear forerunner of a modern approach to a vast array of special applications in audio and motor control.

12.8.1 A Fixed-Gain IC Power Amplifier

Our first example is the LM380 (a product of National Semiconductor Corporation), which is a fixed-gain monolithic power amplifier. A simplified version of the internal circuit of the amplifier[2] is shown in Fig. 12.29. The circuit consists of an input differential amplifier utilizing Q_1 and Q_2 as emitter followers for input buffering, and Q_3 and Q_4 as a differential pair with an emitter resistor R_3. The two resistors R_4 and R_5 provide dc paths to ground for the base currents of Q_1 and Q_2, thus enabling the input signal source to be capacitively coupled to either of the two input terminals.

The differential amplifier transistors Q_3 and Q_4 are biased by two separate currents: Q_3 is biased by a current from the dc supply V_S through the diode-connected transistor Q_{10}, and resistor R_1; Q_4 is biased by a dc current from the output terminal through R_2. Under quiescent conditions (i.e., with no input signal applied) the two bias currents will be equal, and the current through and the voltage across R_3 will be zero. For the emitter current of Q_3 we can write

$$I_3 \simeq \frac{V_S - V_{EB10} - V_{EB3} - V_{EB1}}{R_1}$$

where we have neglected the small dc voltage drop across R_4. Assuming, for simplicity, all V_{EB} to be equal,

$$I_3 \simeq \frac{V_S - 3V_{EB}}{R_1} \tag{12.69}$$

For the emitter current of Q_4 we have

$$I_4 = \frac{V_O - V_{EB4} - V_{EB2}}{R_2} \simeq \frac{V_O - 2V_{EB}}{R_2} \tag{12.70}$$

[2]The main objective of showing this circuit is to point out some interesting design features. The circuit is *not* a detailed schematic diagram of what is actually on the chip.

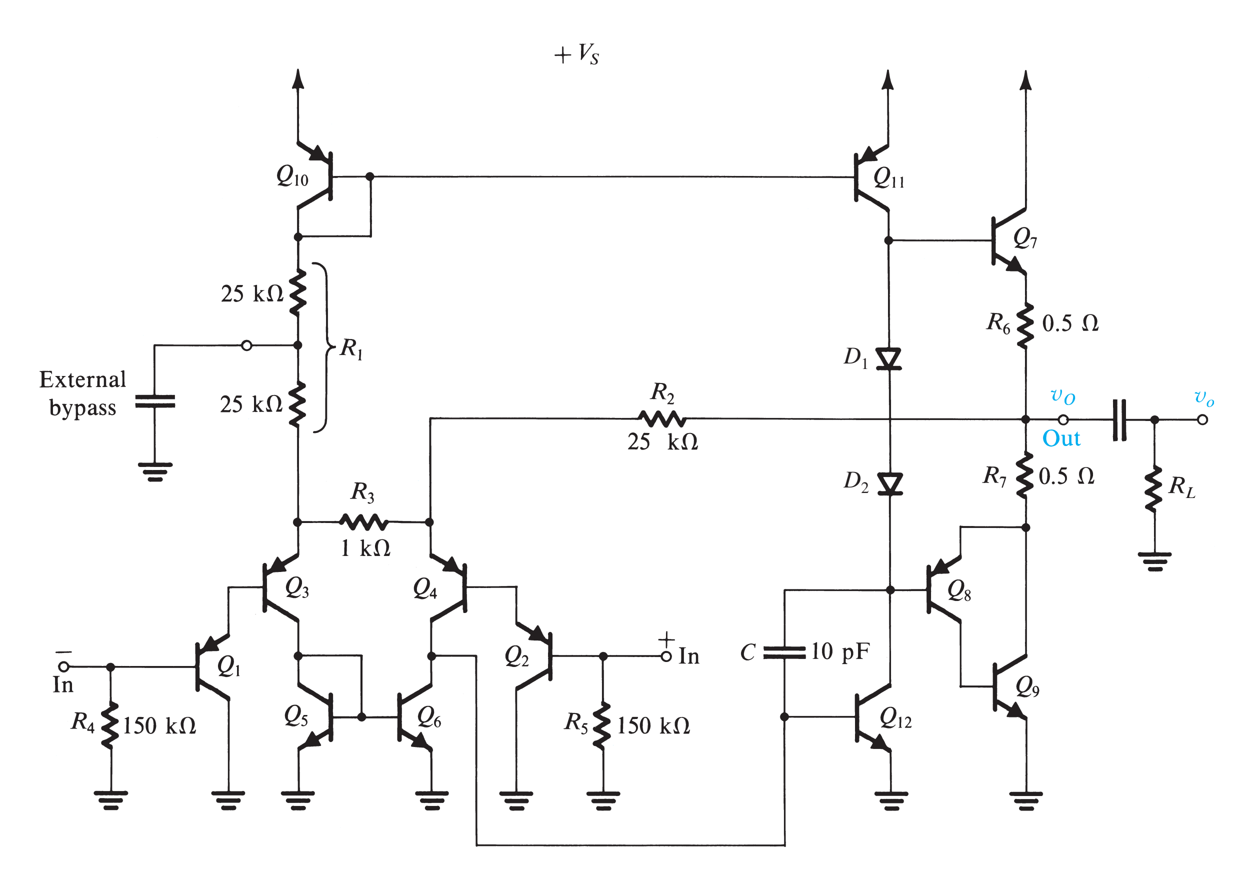


Figure 12.29 The simplified internal circuit of the LM380 IC power amplifier. (Courtesy National Semiconductor Corporation.)

where V_O is the dc voltage at the output, and we have neglected the small drop across R_5. Equating I_3 and I_4 and using the fact that $R_1 = 2R_2$ results in

$$V_O = \frac{1}{2}V_S + \frac{1}{2}V_{EB} \tag{12.71}$$

Thus the output is biased at approximately half the power-supply voltage, as desired for maximum output voltage swing. An important feature is the dc feedback from the output to the emitter of Q_4, through R_2. This dc feedback acts to stabilize the output dc bias voltage at the value in Eq. (12.71). Qualitatively, the dc feedback functions as follows: If for some reason V_O increases, a corresponding current increment will flow through R_2 and into the emitter of Q_4. Thus the collector current of Q_4 increases, resulting in a positive increment in the voltage at the base of Q_{12}. This, in turn, causes the collector current of Q_{12} to increase, thus bringing down the voltage at the base of Q_8 and hence V_O.

Continuing with the description of the circuit in Fig. 12.29, we observe that the differential amplifier (Q_3, Q_4) has a current-mirror load composed of Q_5 and Q_6 (refer to Chapter 9, Section 9.5). The single-ended output voltage signal of the first stage appears at the collector of Q_6 and thus is applied to the base of the second-stage common-emitter amplifier Q_{12}. Transistor Q_{12} is biased by the constant-current source Q_{11}, which also acts as its active load. In actual operation, however, the load of Q_{12} will be dominated by the reflected resistance due to R_L. Capacitor C provides frequency compensation (see Chapter 11).

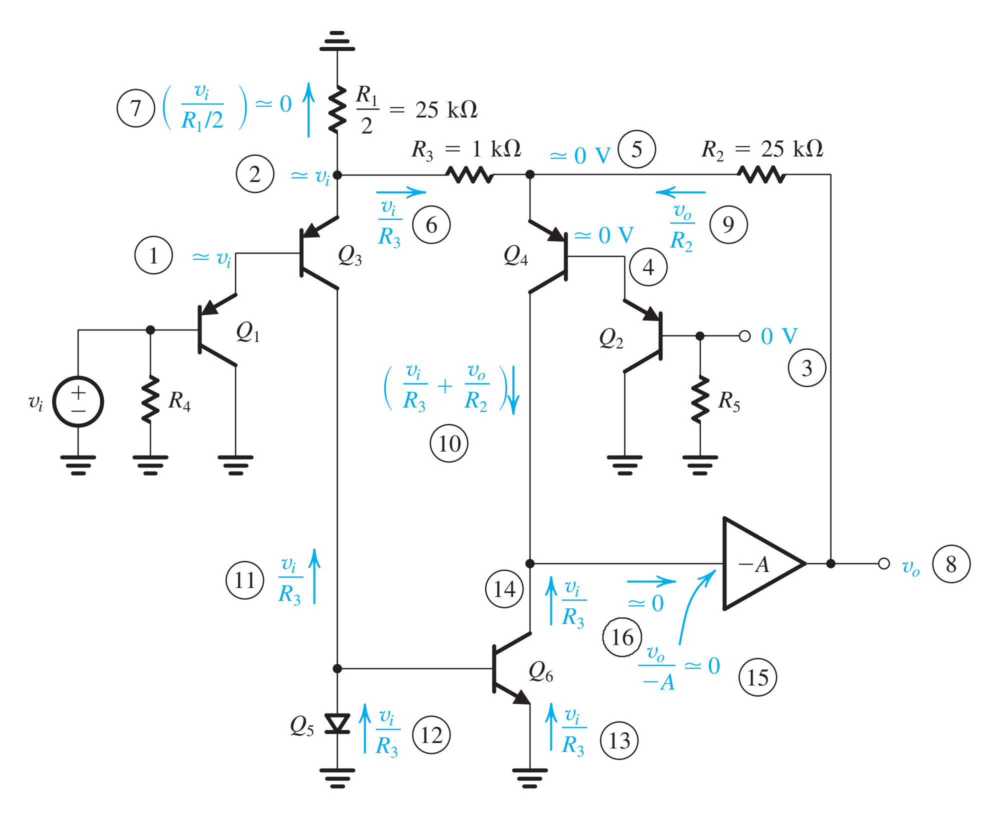


Figure 12.30 Small-signal analysis of the circuit in Fig. 12.29. The circled numbers indicate the order of the analysis steps.

The output stage is class AB, utilizing a compound *pnp* transistor (Q_8 and Q_9). Negative feedback is applied from the output to the emitter of Q_4 via resistor R_2. To find the closed-loop gain consider the small-signal equivalent circuit shown in Fig. 12.30. Here, we have replaced the second-stage common-emitter amplifier and the output stage with an inverting amplifier block with gain A. We shall assume that the amplifier A has high gain and high input resistance, and thus the input signal current into A is negligibly small. Under this assumption, Fig. 12.30 shows the analysis details with an input signal v_i applied to the inverting input terminal. The order of the analysis steps is indicated by the circled numbers. Note that since the input differential amplifier has a relatively large resistance, R_3, in the emitter circuit, most of the applied input voltage appears across R_3. In other words, the signal voltages across the emitter–base junctions of Q_1, Q_2, Q_3, and Q_4 are small in comparison to the voltage across R_3. Accordingly, the voltage gain can be found by writing a node equation at the collector of Q_6:

$$\frac{v_i}{R_3} + \frac{v_o}{R_2} + \frac{v_i}{R_3} = 0$$

which yields

$$\frac{v_o}{v_i} = -\frac{2R_2}{R_3} \simeq -50 \text{ V/V}$$

EXERCISE

12.15 Denoting the total resistance between the collector of Q_6 and ground by R, show, using Fig. 12.30, that

$$\frac{v_o}{v_i} = \frac{-2R_2/R_3}{1+(R_2/AR)}$$

which reduces to $(-2R_2/R_3)$ under the condition that $AR \gg R_2$.

As was demonstrated in Chapter 11, one of the advantages of negative feedback is the reduction of nonlinear distortion. This is the case in the circuit of the LM380.

The LM380 is designed to operate from a single supply V_S in the range of 12 V to 22 V. The selection of supply voltage depends on the value of R_L and the required output power P_L. The manufacturer supplies curves for the device power dissipation versus output power for a given load resistance and various supply voltages. One such set of curves for $R_L = 8\ \Omega$ is shown in Fig. 12.31. Note the similarity to the class B power dissipation curve of Fig. 12.8. In fact, the reader can easily verify that the location and value of the peaks of the curves in Fig. 12.31 are accurately predicted by Eqs. (12.20) and (12.21), respectively (where $V_{CC} = \frac{1}{2}V_S$). The line labeled "3% distortion level" in Fig. 12.31 is the locus of the points on the various curves at

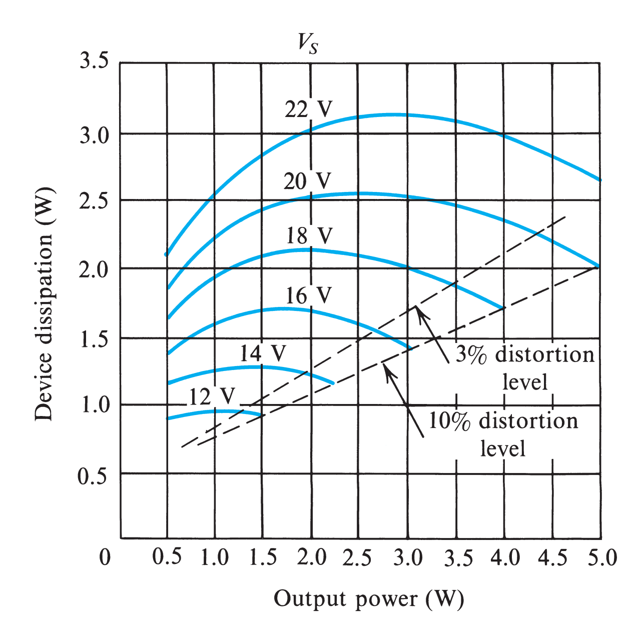


Figure 12.31 Power dissipation (P_D) versus output power (P_L) for the LM380 with $R_L = 8\ \Omega$. (Courtesy National Semiconductor Corporation.)

which the distortion (THD) reaches 3%. A THD of 3% represents the onset of peak clipping due to output-transistor saturation.

EXERCISE

D12.16 It is required to use the LM380 to drive an 8-Ω loudspeaker. Use the curves of Fig. 12.31 to determine the maximum power supply possible while limiting the maximum power dissipation to 2.9 W. If for this application a 3% THD is allowed, find P_L and the peak-to-peak output voltage.
Ans. 20 V; 4.2 W; 16.4 V

12.8.2 The Bridge Amplifier

We conclude this section with a discussion of a circuit configuration that is popular in high-power applications. This is the bridge amplifier configuration shown in Fig. 12.32 utilizing two power op amps, A_1 and A_2. While A_1 is connected in the noninverting configuration with a gain $K = 1 + (R_2/R_1)$, A_2 is connected as an inverting amplifier with a gain of equal magnitude $K = R_4/R_3$. The load R_L is floating and is connected between the output terminals of the two op amps.

If v_I is a sinusoid with amplitude $\hat{V}_i$, the voltage swing at the output of each op amp will be $\pm K\hat{V}_i$, and that across the load will be $\pm 2K\hat{V}_i$. Thus, with op amps operated from ±15-V supplies and capable of providing, say, a ±12-V output swing, an output swing of ±24 V can be obtained across the load of the bridge amplifier.

In designing bridge amplifiers, note should be taken of the fact that the peak current drawn from each op amp is $2K\hat{V}_i/R_L$. This effect can be taken into account by considering the load seen by each op amp (to ground) to be $R_L/2$.

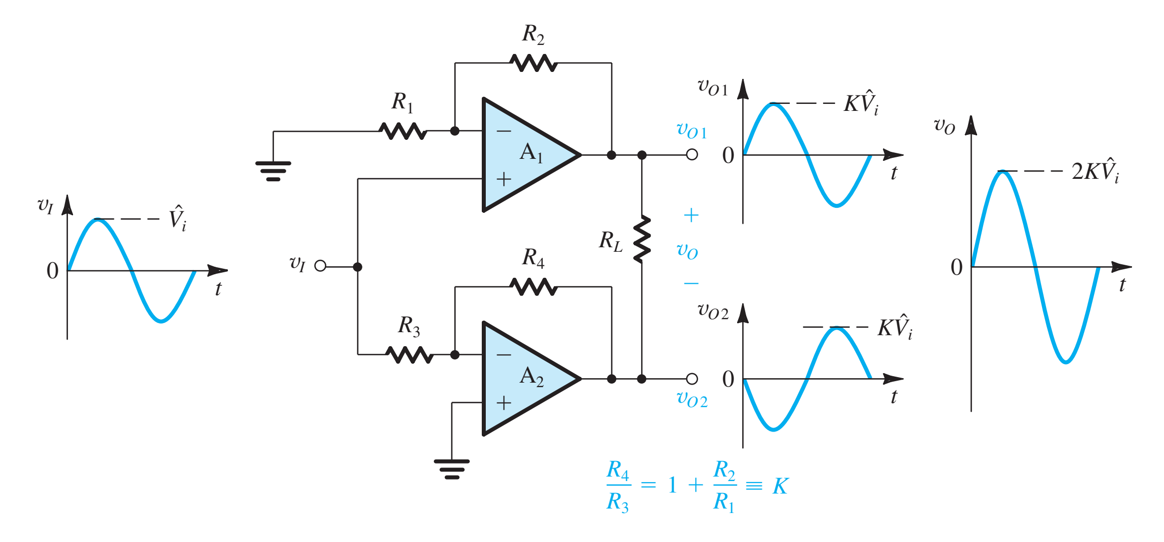


Figure 12.32 The bridge amplifier configuration.

EXERCISE

12.17 Consider the circuit of Fig. 12.32 with $R_1 = R_3 = 10\ \text{k}\Omega$, $R_2 = 5\ \text{k}\Omega$, $R_4 = 15\ \text{k}\Omega$, and $R_L = 8\ \Omega$. Find the voltage gain and the input resistance. The power supply used is ± 18 V. If v_I is a 20-V peak-to-peak sine wave, what is the peak-to-peak output voltage? What is the peak load current? What is the load power?
Ans. 3 V/V; 10 kΩ; 60 V; 3.75 A; 56.25 W

12.9 Class D Power Amplifiers



The class A, B, and AB output stages studied in the previous sections are linear amplifiers. They operate basically as voltage followers, reproducing across the load a replica of the input signal. The output transistors in these amplifiers can dissipate considerable power. This occurs because when providing a large current to the load, the voltage across the transistors can be relatively large. The result is a relatively low power-conversion efficiency, ranging from a theoretical maximum of 25% for the class A stage to 78.5% for the class B and AB stages. In practice, only 50% to 60% efficiencies are achieved in the class B and AB stages.

A very different approach to power amplification, utilized frequently in audio amplifier design, is found in the class D stage. Here, the power dissipated in the output transistors is reduced, theoretically to zero, by operating the transistors as on-off switches.[3] For this to be possible, the audio-frequency signal is converted to a pulse format and the pulses are utilized to operate the output transistor switches. Commonly, pulsewidth modulation (PWM) is utilized. The PWM signal has a frequency at least 10 times higher than the highest audio frequency signal to be amplified. While the amplitude and frequency of the PWM signal remain constant, the magnitude of the audio signal is encoded in the width of the pulses. Specifically, the width of successive pulses is made proportional to the corresponding instantaneous magnitude of the audio signal v_A.

Figure 12.33(a) shows how a PWM signal can be generated. The magnitude of the audio signal v_A is compared to that of a triangular wave v_T whose frequency in the example shown is 10 times that of v_A. The comparison is performed by a circuit building block known as a **comparator**, shown in Fig. 12.33(c) as a black box. We shall study comparators in Chapter 18. Observe that whenever the magnitude of v_A exceeds that of v_T, the comparator output will be at its high level, V^+. Conversely, when the magnitude of v_A falls below that of v_T, the comparator output goes to its low level, V^-. Thus, at the output of the comparator we obtain the pulse waveform shown in Fig. 12.33(b). This waveform has the same frequency f_s of the triangular wave and has standard high and low levels determined by the comparator design. The width of each pulse, t_P, and hence the duty ratio (t_P/T), where $T = 1/f_s$, is proportional to the corresponding instantaneous value of v_A. This is the PWM signal.

[3] An ideal switch has a zero "on" resistance and thus dissipates no power when it is closed and delivering the high load current. Also, it has infinite off resistance and thus dissipates zero power when it is open.

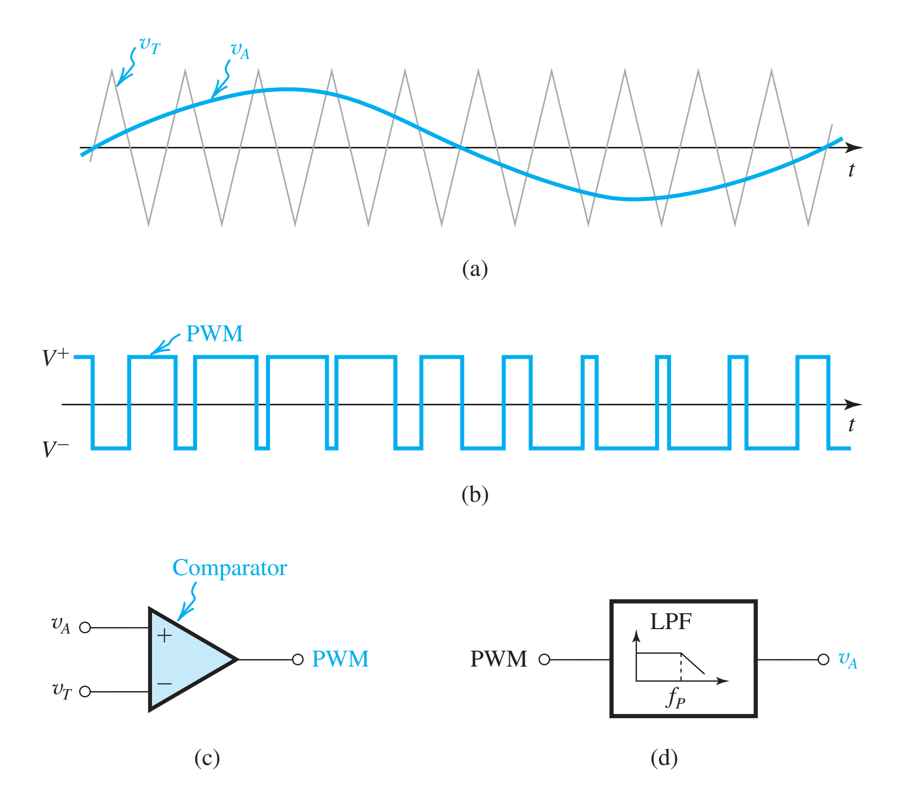


Figure 12.33 (**a**) By comparing the magnitude of the audio signal v_A to that of a triangular wave v_T, the PWM signal in (**b**) can be generated by using the comparator in (**c**). (**d**) The original signal v_A can be recovered from the PWM signal by means of a low-pass filter with a passband frequency f_P slightly larger than the highest audio-frequency component of v_A.

Since the average of a pulse waveform is determined by its duty ratio, the original audio signal v_A can be recovered from the PWM signal by taking the time average of the latter. This in turn can be achieved by passing the PWM signal through a low-pass filter (LPF) whose cutoff frequency is just above the highest frequency of the audio signal.[4] This is shown in block-diagram form in Fig. 12.33(d). The design of low-pass filters will be studied in Chapter 17.

HANS CAMENZIND—THE INVENTOR OF THE CLASS D AMPLIFIER:

In 1966, while working for Mallory (now Duracell), Swiss-born engineer Hans Camenzind (1934–2012) filed a patent on "Pulse-Width Modulation Circuits," introducing the two-state or class D amplifier. Camenzind had begun his U.S. career in 1960, at Transitron, an early semiconductor manufacturer. Later, in 1969, while at Mallory, he filed a second patent on a "Two-State Amplifier." The class D amplifier has revolutionized the field of high-efficiency compact amplifiers. Today these devices, ubiquitous in mobile phones, also see application at enormous power levels for live-performance sound systems with thousands of watts per channel.

[4]The PWM signal has frequency components at f_s and its harmonics, all of which will be higher than the frequency of the audio signal and thus can be easily removed by the low-pass filter.

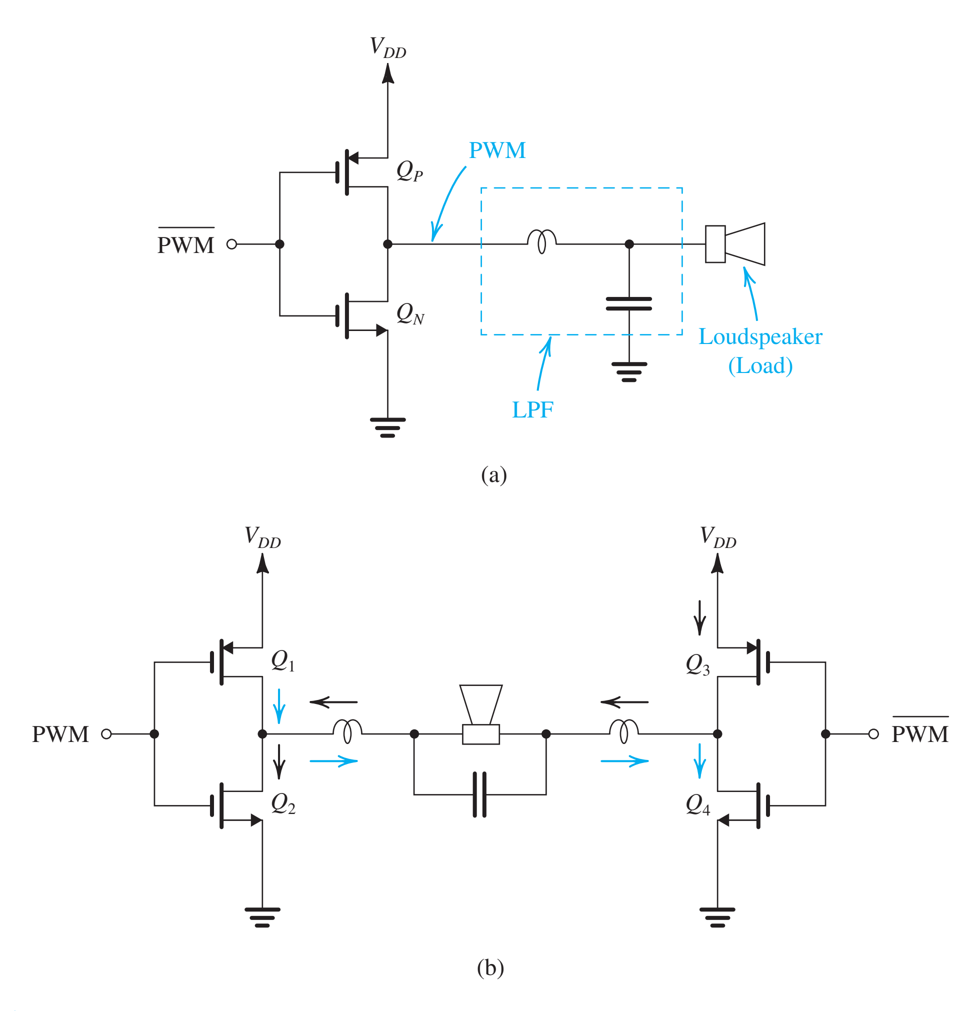


Figure 12.34 Two schemes for driving the load of a class D amplifier. The differential scheme in (**b**) results in doubling the voltage excursion across the load.

Having obtained a pulse waveform in which the audio signal is encoded, we now show how the PWM signal can be used to drive the switches that supply the load power. Two alternative schemes for accomplishing this task are shown in Fig. 12.34. In Fig. 12.34(a), the logical inverse of PWM, denoted $\overline{\text{PWM}}$ and obtained from the comparator by simply exchanging the terminals to which v_A and v_T are applied, is used to drive two complementary MOS switches Q_P and Q_N. These switches connect the output node alternatively to V_{DD} and ground, in effect producing a high-power version of PWM at their drain node. This is the signal applied to the load (shown as a loudspeaker) through a low-pass filter. It follows that v_A appears across the load and the large current required by the low-resistance load is supplied by Q_P and Q_N.

To double the voltage excursion across the load, the scheme in Fig. 12.34(b) can be utilized. Here both PWM and its logical inverse $\overline{\text{PWM}}$ are used in a differential driving arrangement. When PWM is high and thus $\overline{\text{PWM}}$ is low, Q_3 and Q_2 are turned on while Q_1 and Q_4 are off. Thus current flows from V_{DD} to ground through the load (from right to left). The opposite happens when PWM is low. Thus the voltage across the load will be twice that obtained with

the arrangement in Fig. 12.34(a). The circuit of the differential driving arrangement is known as an H bridge[5] and can result in a maximum sinusoidal output voltage of amplitude V_{DD}.

The description above leads to the conclusion that the power-conversion efficiency of a class D amplifier is 100%. This, of course, is only a theoretical limit. In practice, the power transistors exhibit finite on-resistances that lead to conduction losses. Also, every time the power MOSFETs are turned on and off, the gate and load capacitances are charged and discharged, resulting in power loss in the driving circuit. In addition, due to the finite switching speed of power MOSFETs, there will be a momentary short between V_{DD} and ground during every switching cycle. The resulting *shoot-through current* is exactly the same as that encountered in a CMOS inverter (see Chapter 14), but it could be of much larger magnitude owing to the large transistor size. All these sources of power dissipation cause the power-conversion efficiency to be in the 85% to 95% range, still much larger than is achieved in a class AB stage.

As a final note, distortion in the class D amplifier can be contributed by the PWM modulation scheme, by inaccuracy of the duty ratio at the output node due to finite switching speeds of the power MOSFETs, and by the quality factor of the L and C elements in the output filter. Class D amplifiers typically exhibit THD of 0.1% to 1% at best. Another imperfection of the class D amplifiers is the generation of unwanted switching noise, usually in the inaudible range as electromagnetic interference. As a result, class D amplifiers are most useful in applications where power-conversion efficiency is of paramount importance. Class A and AB amplifiers can achieve THD figures of less than 0.01% and are mostly used in high-fidelity applications.

EXERCISES

12.18 Consider the comparator in Fig. 12.33(c) with the triangular wave v_T having ± 10 V peak voltages and comparator output levels of ± 10 V. Find the duty ratio D and the average of the output voltage for the case in which v_A is a constant voltage of magnitude (a) 0 V; (b) +5 V; (c) +10 V; (d) −5 V; (e) −10 V

Ans. (a) 50%, 0 V; (b) 75%, +5 V; (c) 100%, +10 V; (d) 25%, −5 V; (e) 0 %; −10 V

12.19 If the audio signal v_A has a frequency spectrum of 20 Hz to 20 kHz, what is an appropriate value for f_s? Now if the low-pass filter is of second order, with its passband edge at 20 kHz and its gain falling off at 40 dB/decade, what is the attenuation encountered by the PWM component with frequency f_s?

Ans. 200 kHz; 40 dB

12.20 If the differential switching scheme shown in Fig. 12.34(b) is utilized and v_A is a sine wave, what is the maximum peak amplitude achieved across R_L and what is the maximum power delivered to R_L? Evaluate these quantities for $V_{DD} = 35$ V and $R_L = 8\ \Omega$. Now, if the power-conversion efficiency is 90%, what is the power delivered by the power supplies?

Ans. V_{DD}, $\dfrac{V_{DD}^2}{2R_L}$; 35 V, 76.6 W; 85.1 W

[5]The name H arises from the resemblance of the circuit diagram to the letter H: The switches represent the vertical strokes of H and the load with the filter represents the horizontal stroke.

12.10 Power Transistors



The BJTs and MOSFETs that are utilized in the design of the output stages and power amplifiers studied in this chapter can be called upon to conduct currents in the ampere range, to support voltages in excess of 100 V, and to withstand power dissipation in the tens-of-watts range. Hence, they are called power transistors. In this section, we study the characteristics, specifications, and thermal operation of power transistors.[6]

12.10.1 Packages and Heat Sinks

Power transistors are basically larger versions of their small-signal counterparts; hence they retain similar characteristics. However, as will be explained shortly, their structures are modified for optimal voltage and current capabilities. Also, discrete power transistors are housed in special packages such as those shown in Fig. 12.35. The packages are usually mounted on **heat sinks**, special metal surfaces whose function is to facilitate the conduction of heat away from the transistor, thus keeping its internal temperature within safe operating limits. We shall have more to say about thermal issues in Section 12.10.4. A typical heat sink is shown in Fig. 12.36.

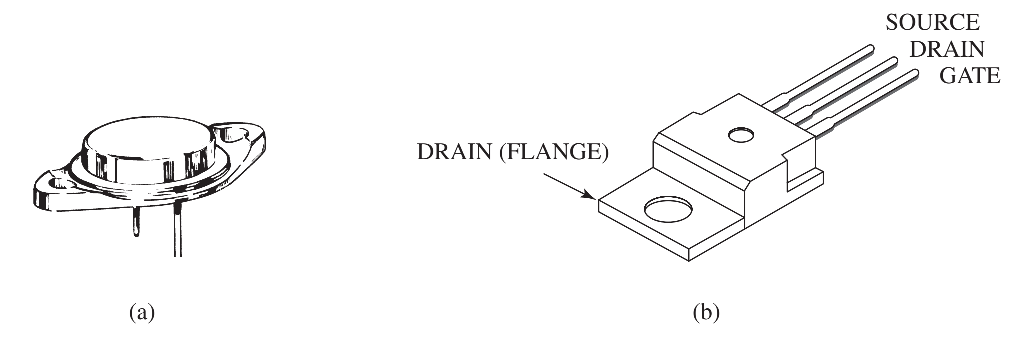


Figure 12.35 Most popular packages for power transistors: (**a**) TO-03 metal package; (**b**) TO-220 plastic package.

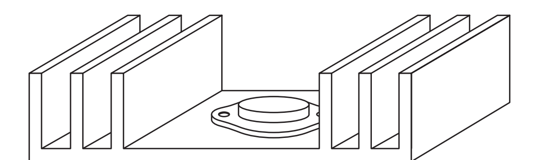


Figure 12.36 Typical heat sink.

[6]Other semiconductor devices utilized in power electronic applications include thyristors and silicon-controlled rectifiers (SCRs). These are usually used in applications requiring much higher current and voltage ratings than those studied in this chapter. As well, a more recent device, the insulated-gate bipolar transistor, or IGBT, has a merged MOS/bipolar structure and combines the advantages of both. It is used in very high current applications and is not studied in this book.

12.10.2 Power BJTs

Device Structure The power BJT utilizes a variation on the basic structure in Fig. 6.7. Specifically:

1. To increase the current-handling capability of the BJT while maintaining the current density at a reasonable level, the emitter area is made much larger. This is accomplished by utilizing multiple emitter regions (called "emitter fingers") and connecting them together, as shown in the simplified device cross section in Fig. 12.37. To reduce the extrinsic base resistance r_x, the width of each emitter finger is kept small.
2. To support higher voltages without device breakdown, the base is made wider (with the attendant reduction in β), and the collector is made thicker and its doping lighter. Again, these features are indicated in Fig. 12.37.

Device Parameters As a result of the structural differences, the parameters of power BJTs can differ somewhat from those of small-signal devices. Important differences include the following.

1. The current gain β is low, typically in the range of 10 to 80, but can be as low as 5. Here it is important to recall that β is a function of current and has a positive temperature coefficient (refer to Fig. 6.34).
2. The maximum collector current $I_{C\max}$ is typically in the ampere range but can be as high as 100 A.
3. The breakdown voltage (BV_{CEO}; refer to Fig. 6.33) is typically 50 V to 100 V but can be as high as 500 V.
4. I_{CBO} is large (a few tens of microamps) and, as usual, doubles for every 10°C rise in temperature.

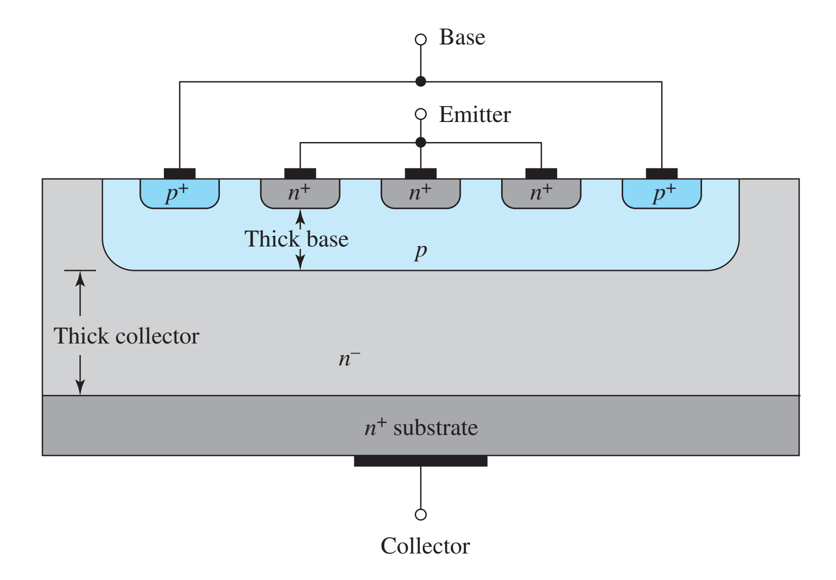


Figure 12.37 Cross section of a power BJT.

5. At high currents, r_π becomes small (a few ohms) and the extrinsic base resistance r_x becomes important.
6. The transition frequency f_T is low (a few megahertz), C_μ is large (hundreds of picofarads), and C_π is even larger.
7. At high currents, the exponential i_C–v_{BE} relationship exhibits a factor-of-2 reduction in the exponent: that is, $i_C = I_S e^{v_{BE}/2V_T}$.

At large collector currents, the low β results in the requirement for a large base current. This can complicate the design of the circuit that drives the output transistors. The Darlington configuration discussed in Section 12.6 can be employed to provide a higher effective β.

The BJT Safe Operating Area The power dissipated in a BJT results in an increase in its temperature and thus power dissipation must be limited. The manufacturer specifies the maximum power that can be safely dissipated. In addition, the manufacturer provides a plot of the safe operating area (SOA) in the i_C–v_{CE} plane. The SOA specification takes the form shown in Fig. 12.38; the following paragraph numbers correspond to the boundaries on the sketch.

1. The maximum allowable current $I_{C\max}$. Exceeding this current on a continuous basis can result in melting the wires that bond the device to the package terminals.
2. The maximum power dissipation hyberbola. This is the locus of the points for which $v_{CE} i_C = P_{D\max}$. The specified $P_{D\max}$ corresponds to a specific temperature of the transistor case, T_{C0}. If the case temperature T_C is higher than T_{C0}, a lower value of $P_{D\max}$ and a correspondingly lower hyberbola apply. This point will be explained in Section 12.10.4. Although the operating point can be allowed to move temporarily above the hyberbola, the *average* power dissipation should not be allowed to exceed the applicable $P_{D\max}$.
3. The **second-breakdown** limit. Second breakdown is a phenomenon that results because current flow across the emitter–base junction is not uniform. Rather, the current density is greatest near the periphery of the junction. This *"current-crowding"* gives rise to increased localized power dissipation and hence temperature rise (at

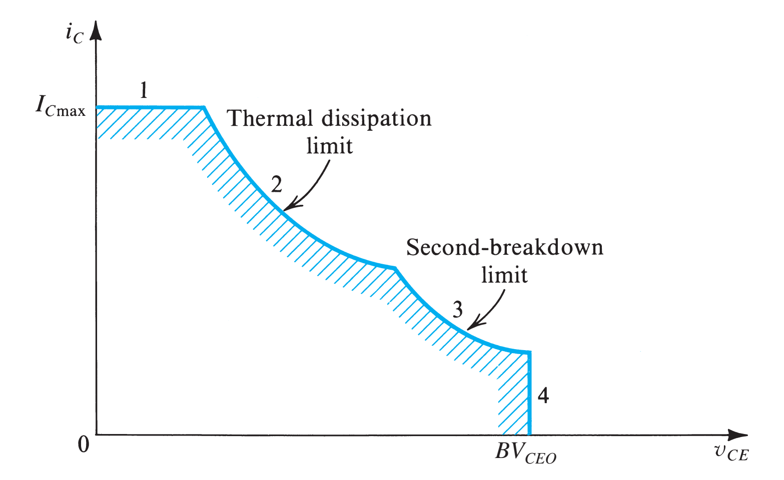


Figure 12.38 Safe operating area (SOA) of a BJT.

locations called **hot spots**). Since a temperature rise causes an increase in current, a localized form of **thermal runaway** can occur, leading to junction destruction.

4. The collector-to-emitter breakdown voltage BV_{CEO}. The instantaneous value of v_{CE} should never be allowed to exceed BV_{CEO}; otherwise, avalanche breakdown of the collector–base junction may occur (Section 6.4).

Finally, it should be noted that logarithmic scales are usually used for i_C and v_{CE} leading to a safe-operating-area boundary that consists of straight lines.

12.10.3 Power MOSFETs

Power MOSFETs have in the past number of years gained popularity in the design of power electronic circuits. This is a result of the following properties.

1. Unlike BJTs, MOSFETs do not require dc gate drive current. This greatly simplifies the design of the driving circuitry.
2. MOSFETs can operate at much higher switching speeds than BJTs, a definite advantage for power circuits employing switching, such as class D amplifiers.
3. MOSFETs do not suffer from secondary breakdown, thus benefiting from an extension of SOA.
4. The thermal characteristics of the MOSFET, as we shall see shortly, are superior to those of the BJT.

Structure of the Power MOSFET The MOSFET structure studied in Chapter 5 (Fig. 5.1) is not suitable for high-power applications. To appreciate this fact, recall that the drain current of an *n*-channel MOSFET operating in the saturation region is given by

$$i_D = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)(v_{GS}-V_t)^2 \tag{12.72}$$

It follows that to increase the current capability of the MOSFET, its width W should be made large and its channel length L should be made as small as possible. Unfortunately, however, reducing the channel length of the standard MOSFET structure results in a drastic reduction in its breakdown voltage. Specifically, the depletion region of the reverse-biased body-to-drain junction spreads into the short channel, resulting in breakdown at a relatively low voltage. Thus the resulting device would not be capable of handling the high voltages typical of power-transistor applications. For this reason, new structures had to be found for fabricating short-channel (1- to 2-μm) MOSFETs with high breakdown voltages.

At the present time the most popular structure for a power MOSFET is the **double-diffused** or **DMOS transistor** shown in Fig. 12.39. As indicated, the device is fabricated on a lightly doped *n*-type substrate with a heavily doped region at the bottom for the drain contact. Two diffusions[7] are employed, one to form the *p*-type body region and another to form the *n*-type source region.

The DMOS device operates as follows. Application of a positive gate voltage, v_{GS}, greater than the threshold voltage V_t, induces a lateral *n* channel in the *p*-type body region underneath the gate oxide. The resulting channel is short; its length is denoted L in Fig. 12.39. Current is then conducted by electrons from the source moving through the resulting short channel

[7] See Appendix A for a description of the IC fabrication process.

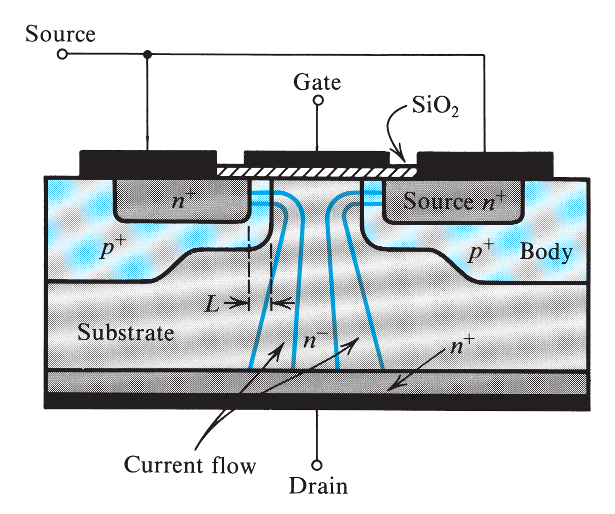


Figure 12.39 Double-diffused vertical MOS transistor (DMOS).

to the substrate and then vertically down the substrate to the drain. This should be contrasted with the lateral current flow in the standard small-signal MOSFET structure (Chapter 5).

Even though the DMOS transistor has a short channel, its breakdown voltage can be very high (as high as 600 V). This is because the depletion region between the substrate and the body extends mostly in the lightly doped substrate and does not spread into the channel. The result is a MOS transistor that simultaneously has a high current capability (50 A is possible) as well as the high breakdown voltage just mentioned. Finally, we note that the vertical structure of the device provides efficient utilization of the silicon area.

Many unit devices such as that in Fig. 12.39 are usually fabricated on a chip and connected in parallel to achieve the required high current capability. Various layout and packing arrangements have been utilized, including an efficient design utilizing hexagons, termed HEXFET and available from International Rectifier.

Characteristics of Power MOSFETs In spite of their radically different structure, power MOSFETs exhibit characteristics that are quite similar to those of the small-signal MOSFETs studied in Chapter 5. Important differences exist, however, and these are discussed next.

Power MOSFETs have threshold voltages in the range of 1 V to 4 V. In saturation, the drain current is related to v_{GS} by the square-law characteristic of Eq. (12.72). However, the i_D–v_{GS} characteristic becomes linear for larger values of v_{GS}. The linear portion of the characteristic occurs as a result of the high electric field along the short channel, causing the velocity of charge carriers to reach an upper limit, a phenomenon known as **velocity saturation**.[8] The linear i_D–v_{GS} relationship implies a constant g_m in the velocity-saturation region.

Of considerable interest in the design of MOS power circuits is the variation of the MOSFET characteristics with temperature, illustrated in Fig. 12.40. Observe that there is a value of v_{GS} (in the range of 4 V to 6 V for most power MOSFETs) at which the temperature coefficient of i_D is zero. At higher values of v_{GS}, i_D exhibits a negative temperature coefficient. This is a significant property: It implies that a MOSFET operating beyond the zero-temperature-coefficient point does not risk the possibility of thermal runaway. This is *not* the case, however, at low currents (i.e., lower than the zero-temperature-coefficient point). In

[8]Velocity saturation occurs also in standard MOSFET structures when the channel length is in the submicron range. We shall discuss velocity saturation in some detail in Section 15.1.

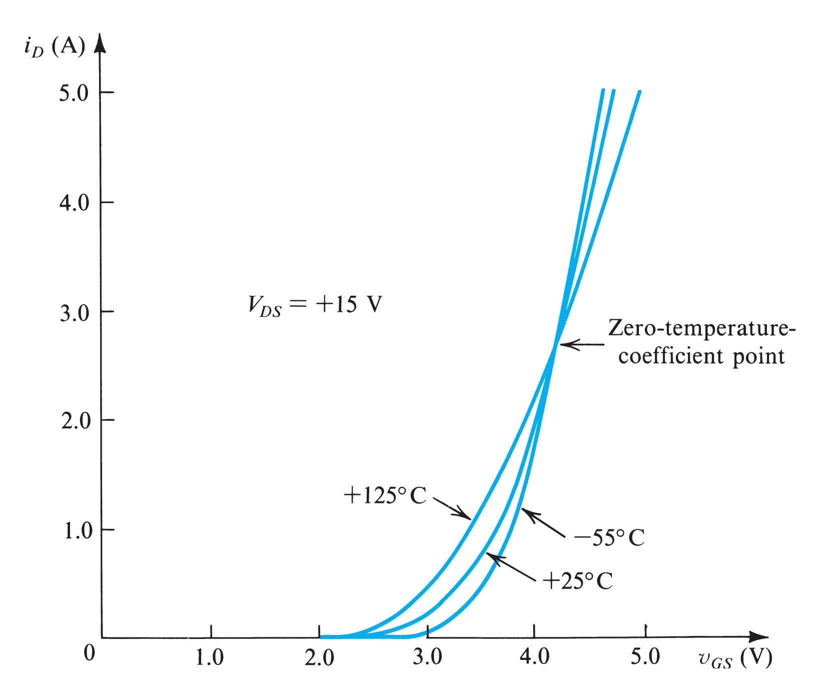


Figure 12.40 The i_D–v_{GS} characteristic curve of a power MOS transistor (IRF 630, Siliconix) at case temperatures of –55°C, +25°C, and +125°C. (Courtesy of Siliconix Inc.)

the (relatively) low-current region, the temperature coefficient of i_D is positive, and the power MOSFET can easily suffer thermal runaway (with unhappy consequences). Since class AB output stages are biased at low currents, means must be provided to guard against thermal runaway.

The reason for the positive temperature coefficient of i_D at low currents is that $v_{OV} = (v_{GS} - V_t)$ is relatively low, and the temperature dependence is dominated by the negative temperature coefficient of V_t (in the range of –3 mV/°C to –6 mV/°C), which causes v_{OV} to rise with temperature.

12.10.4 Thermal Considerations

Power transistors dissipate large amounts of power. The dissipated power is converted into heat, which raises the junction temperature. However, the junction temperature T_J must not be allowed to exceed a specified maximum, $T_{J\max}$; otherwise the transistor could suffer permanent damage. For silicon devices, $T_{J\max}$ is in the range of 150°C to 200°C.

Thermal Resistance Consider first the situation of a transistor operating in free air—that is, with no special arrangements for cooling. The heat dissipated in the transistor will be conducted away from the junction to the transistor case, and from the case to the surrounding environment. In a steady state in which the transistor is dissipating P_D watts, the temperature rise of the junction relative to the surrounding ambience can be expressed as



$$T_J - T_A = \theta_{JA} P_D \qquad (12.73)$$

where θ_{JA} is the **thermal resistance** between junction and ambience, having the units of degrees Celsius per watt. Note that θ_{JA} simply gives the rise in junction temperature over the ambient temperature for each watt of dissipated power. Since we wish to be able to dissipate large amounts of power without raising the junction temperature above $T_{J\max}$, it is desirable to have, for the thermal resistance θ_{JA}, as small a value as possible. For operation in free air, θ_{JA} depends primarily on the type of case in which the transistor is packaged. The value of θ_{JA} is usually specified on the transistor data sheet.

Equation (12.73), which describes the thermal-conduction process, is analogous to Ohm's law, which describes the electrical-conduction process. In this analogy, power dissipation corresponds to current, temperature difference corresponds to voltage difference, and thermal resistance corresponds to electrical resistance. Thus, we may represent the thermal-conduction process by the electric circuit shown in Fig. 12.41.

Power Dissipation versus Temperature The transistor manufacturer usually specifies the maximum junction temperature $T_{J\max}$, the maximum allowable power dissipation at a particular ambient temperature T_{A0} (usually 25°C), and the thermal resistance θ_{JA}. In addition, a graph such as that shown in Fig. 12.42 is usually provided. The graph simply states that for operation at ambient temperatures below T_{A0}, the device can safely dissipate the rated value of P_{D0} watts. However, if the device is to be operated at higher ambient temperatures, the maximum allowable power dissipation must be **derated** according to the straight line shown in Fig. 12.42. The **power-derating curve** is a graphical representation of Eq. (12.73). Specifically, note that if the ambient temperature is T_{A0} and the power dissipation is at the maximum allowed (P_{D0}), then the junction temperature will be $T_{J\max}$. Substituting these quantities in Eq. (12.73) results in

$$\theta_{JA} = \frac{T_{J\max} - T_{A0}}{P_{D0}} \tag{12.74}$$

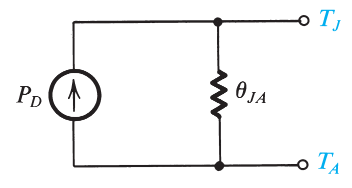


Figure 12.41 Electrical equivalent circuit of the thermal-conduction process; $T_J - T_A = P_D\theta_{JA}$.

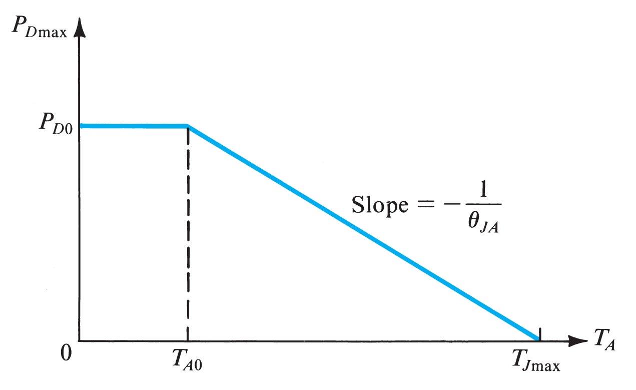


Figure 12.42 Maximum allowable power dissipation versus ambient temperature for a BJT operated in free air. This is known as a "power-derating" curve.

which is the inverse of the slope of the power-derating straight line. At an ambient temperature T_A, higher than T_{A0}, the maximum allowable power dissipation $P_{D\max}$ can be obtained from Eq. (12.73) by substituting $T_J = T_{J\max}$; thus,

$$P_{D\max} = \frac{T_{J\max} - T_A}{\theta_{JA}} \tag{12.75}$$

Observe that as T_A approaches $T_{J\max}$, the allowable power dissipation decreases; the lower thermal gradient limits the amount of heat that can be removed from the junction. In the extreme situation of $T_A = T_{J\max}$, no power can be dissipated because no heat can be removed from the junction.

Example 12.7

A BJT is specified to have a maximum power dissipation P_{D0} of 2 W at an ambient temperature T_{A0} of 25°C, and a maximum junction temperature $T_{J\max}$ of 150°C. Find the following:

(a) the thermal resistance θ_{JA}
(b) the maximum power that can be safely dissipated at an ambient temperature of 50°C
(c) the junction temperature if the device is operating at $T_A = 25°\text{C}$ and is dissipating 1 W

Solution

(a) $$\theta_{JA} = \frac{T_{J\max} - T_{A0}}{P_{D0}} = \frac{150 - 25}{2} = 62.5°\text{C/W}$$

(b) $$P_{D\max} = \frac{T_{J\max} - T_A}{\theta_{JA}} = \frac{150 - 50}{62.5} = 1.6 \text{ W}$$

(c) $$T_J = T_A + \theta_{JA} P_D = 25 + 62.5 \times 1 = 87.5°\text{C}$$

Transistor Case and Heat Sink The thermal resistance between junction and ambience, θ_{JA}, can be expressed as

$$\theta_{JA} = \theta_{JC} + \theta_{CA} \tag{12.76}$$

where θ_{JC} is the thermal resistance between junction and transistor case (package) and θ_{CA} is the thermal resistance between case and ambience. For a given transistor, θ_{JC} is fixed by the device design and packaging. The device manufacturer can reduce θ_{JC} by encapsulating the device in a relatively large metal case, such as that in Fig. 12.35(a), and placing the collector (where most of the heat is dissipated) in direct contact with the case.

Although the circuit designer has no control over θ_{JC} (once a particular transistor has been selected), the designer can considerably reduce θ_{CA} below its free-air value (specified by the manufacturer as part of θ_{JA}). Reduction of θ_{CA} can be effected by providing means to facilitate heat transfer from case to ambience. A popular approach is to bolt the transistor to

the chassis or to an extended metal surface such as the heat sink shown in Fig. 12.36. Heat is easily conducted from the transistor case to the heat sink; that is, the thermal resistance θ_{CS} is usually very small. Also, heat is efficiently transferred (by convection and radiation) from the heat sink to the ambience, resulting in a low thermal resistance θ_{SA}. Thus, if a heat sink is utilized, the case-to-ambience thermal resistance given by

$$\theta_{CA} = \theta_{CS} + \theta_{SA} \tag{12.77}$$

can be small because its two components can be made small by the choice of an appropriate heat sink. For example, in very high-power applications the heat sink is usually equipped with fins that further facilitate cooling by radiation and convection.

The electrical analog of the thermal-conduction process when a heat sink is employed is shown in Fig. 12.43, from which we can write

$$T_J - T_A = P_D(\theta_{JC} + \theta_{CS} + \theta_{SA}) \tag{12.78}$$

As well as specifying θ_{JC}, the device manufacturer usually supplies a derating curve for $P_{D\max}$ versus the case temperature, T_C. Such a curve is shown in Fig. 12.44. Note that the slope of the power-derating straight line is $-1/\theta_{JC}$. For a given transistor, the maximum power dissipation at a *case temperature* T_{C0} (usually 25°C) is much greater than that at an *ambient temperature*

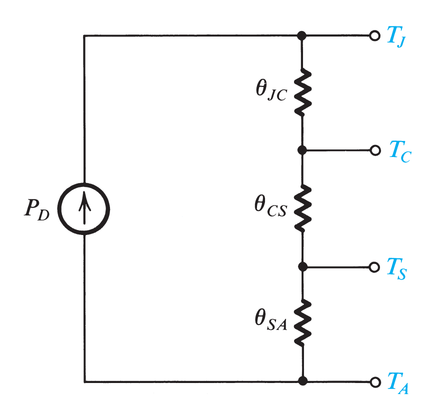


Figure 12.43 Electrical analog of the thermal-conduction process when a heat sink is utilized.

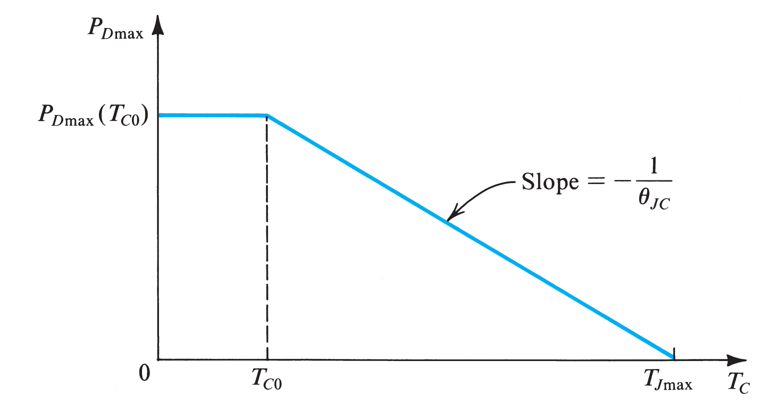


Figure 12.44 Maximum allowable power dissipation versus transistor-case temperature.

T_{A0} (usually 25°C, because $\theta_{JC} \ll \theta_{JA}$). If the device can be maintained at a case temperature T_C, $T_{C0} \leq T_C \leq T_{J\max}$, then the maximum safe power dissipation is obtained when $T_J = T_{J\max}$,

$$P_{D\max} = \frac{T_{J\max} - T_C}{\theta_{JC}} \tag{12.79}$$

Example 12.8

A BJT is specified to have $T_{J\max} = 150$°C and to be capable of dissipating maximum power as follows:

$$40 \text{ W at } T_C = 25\text{°C}$$
$$2 \text{ W at } T_A = 25\text{°C}$$

Above 25°C, the maximum power dissipation is to be derated linearly with $\theta_{JC} = 3.12$°C/W and $\theta_{JA} = 62.5$°C/W. Find the following:

(a) The maximum power that can be dissipated safely by this transistor when operated in free air at $T_A = 50$°C.
(b) The maximum power that can be dissipated safely by this transistor when operated at an ambient temperature of 50°C, but with a heat sink for which $\theta_{CS} = 0.5$°C/W and $\theta_{SA} = 4$°C/W. Find the temperature of the case and of the heat sink.
(c) The maximum power that can be dissipated safely if an *infinite heat sink* is used and $T_A = 50$°C.

Solution

(a)

$$P_{D\max} = \frac{T_{J\max} - T_A}{\theta_{JA}} = \frac{150 - 50}{62.5} = 1.6 \text{ W}$$

(b) With a heat sink, θ_{JA} becomes

$$\theta_{JA} = \theta_{JC} + \theta_{CS} + \theta_{SA}$$
$$= 3.12 + 0.5 + 4 = 7.62\text{°C/W}$$

Thus,

$$P_{D\max} = \frac{150 - 50}{7.62} = 13.1 \text{ W}$$

Figure 12.45 shows the thermal equivalent circuit with the various temperatures indicated.

(c) An infinite heat sink, if it existed, would cause the case temperature T_C to equal the ambient temperature T_A. The infinite heat sink has $\theta_{CA} = 0$. Obviously, one cannot buy an infinite heat sink; nevertheless, this terminology is used by some manufacturers to describe the power-derating curve of Fig. 12.44.

The abscissa is then labeled T_A and the curve is called "power dissipation versus ambient temperature with an infinite heat sink." For our example, with infinite heat sink,

$$P_{D\max} = \frac{T_{J\max} - T_A}{\theta_{JC}} = \frac{150 - 50}{3.12} = 32 \text{ W}$$

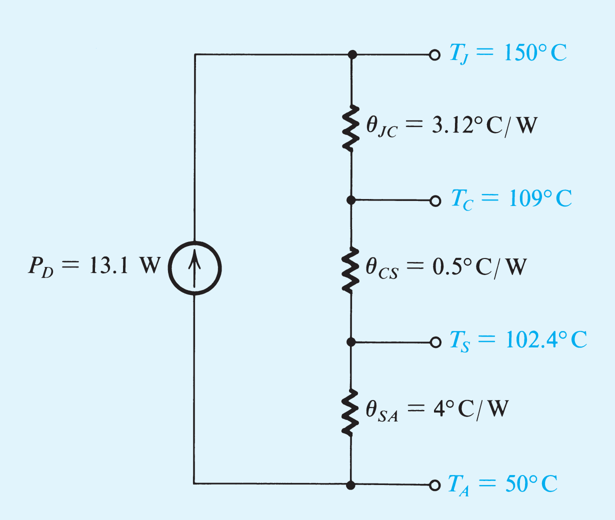


Figure 12.45 Thermal equivalent circuit for Example 12.8.

The advantage of using a heat sink is clearly evident from Example 12.8: With a heat sink, the maximum allowable power dissipation increases from 1.6 W to 13.1 W. Also note that although the transistor considered can be called a "40-W transistor," this level of power dissipation cannot be achieved in practice; that would require an infinite heat sink and an ambient temperature $T_A \leq 25$°C.

EXERCISE

12.21 The 2N6306 power transistor is specified to have $T_{J\max} = 200$°C and $P_{D\max} = 125$ W for $T_C \leq 25$°C. For $T_C \geq 25$°C, $\theta_{JC} = 1.4$°C/W. If in a particular application this device is to dissipate 50 W and operate at an ambient temperature of 25°C, find the maximum thermal resistance of the heat sink that must be used (i.e., θ_{SA}). Assume $\theta_{CS} = 0.6$°C/W. What is the case temperature, T_C?
Ans. 1.5°C/W; 130°C

Summary

- Output stages are classified according to the transistor conduction angle: class A (360°), class AB (slightly more than 180°), class B (180°), and class C (less than 180°).
- The most common class A output stage is the emitter follower. It is biased at a current greater than the peak load current.
- The class A output stage dissipates its maximum power under quiescent conditions ($v_O = 0$). It achieves a maximum power-conversion efficiency of 25%.
- The class B stage is biased at zero current, and thus dissipates no power in quiescence.
- The class B stage can achieve a power-conversion efficiency as high as 78.5%. It dissipates its maximum power for $\hat{V}_o = (2/\pi)V_{CC}$.
- The class B stage suffers from crossover distortion.
- The class AB output stage is biased at a small current; thus both transistors conduct for small input signals, and crossover distortion is virtually eliminated.
- Except for an additional small quiescent power dissipation, the power relationships of the class AB stage are similar to those in class B.
- To guard against the possibility of thermal runaway, the bias voltage of the class AB circuit is made to vary with temperature in the same manner as does V_{BE} of the output transistors.
- Use of the Darlington configuration in the class AB output stage reduces the base-current drive requirement. In integrated circuits, the compound *pnp* configuration is commonly used.
- Output stages are usually equipped with circuitry that, in the event of a short circuit, can turn on and limit the base-current drive, and hence the emitter current, of the output transistors.
- The classical CMOS class AB output stage suffers from reduced output signal swing. This problem can be overcome by replacing the source-follower output transistors with a pair of complementary devices operating in the common-source configuration.
- The CMOS class AB output stage with common-source transistors allows the output voltage to swing to within an overdrive voltage from each of the two power supplies. Utilizing amplifiers in the feedback path of each of the output transistors reduces both the output resistance and the gain error of the stage.
- IC power amplifiers consist of a small-signal voltage amplifier cascaded with a high-power output stage. Overall feedback is applied either on-chip or externally.
- The bridge amplifier configuration provides, across a floating load, a peak-to-peak output voltage that is twice that possible from a single amplifier with a grounded load.
- Class D amplifiers convert the audio signal into a pulsewidth-modulated (PWM) signal. The latter is then used to drive complementary MOS switches that supply the load with power. A low-pass filter is utilized to eliminate the high-frequency components introduced by the switching waveform. Power-conversion efficiences in the range of 85% to 90% are achieved.
- MOSFETs have gained popularity over BJTs in the design of high-power output stages. This is due to their higher speed of operation and to the fact that they do not need a steady supply of gate currents, which allows the use of relatively simple driving circuitry.
- The DMOS transistor is a short-channel power device capable of both high-current and high-voltage operation.
- The drain current of a power MOSFET exhibits a positive temperature coefficient at low currents, and thus the device can suffer thermal runaway. At high currents the temperature coefficient of i_D is negative.
- To facilitate the removal of heat from the silicon chip, power devices are usually mounted on heat sinks. The maximum power that can be safely dissipated in the device is given by

$$P_{D\max} = \frac{T_{J\max} - T_A}{\theta_{JC} + \theta_{CS} + \theta_{SA}}$$

  where $T_{J\max}$ and θ_{JC} are specified by the manufacturer, while θ_{CS} and θ_{SA} depend on the heat-sink design.

Computer Simulation Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as allowable signal swing and amplifier nonlinear distortion. Instructions to assist in setting up PSpice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 12.2: Class A Output Stage

12.1 A class A emitter follower, biased using the circuit shown in Fig. 12.2, uses $V_{CC} = 10$ V, $R = R_L = 1$ kΩ, with all transistors (including Q_3) identical. Assume $V_{BE} = 0.7$ V, $V_{CE\,sat} = 0.3$ V, and β to be very large. For linear operation, what are the upper and lower limits of output voltage, and the corresponding inputs? How do these values change if the emitter–base junction area of Q_3 is made twice as big as that of Q_2? Half as big?

12.2 A source-follower circuit using NMOS transistors is constructed following the pattern shown in Fig. 12.2. All three transistors used are identical, with $V_t = 0.5$ V and $\mu_n C_{ox} W/L = 20$ mA/V^2; $V_{CC} = 2.5$ V, $R = R_L = 1$ kΩ. For linear operation, what are the upper and lower limits of the output voltage, and the corresponding inputs?

D 12.3 Using the follower configuration shown in Fig. 12.2 with ±5-V supplies, provide a design capable of ±3-V outputs with a 1-kΩ load, using the smallest possible total supply current. You are provided with four identical, high-β BJTs and a resistor of your choice. Select a standard resistor value of 5% tolerance, and specify the maximum power drawn from the negative supply.

D 12.4 An emitter follower using the circuit of Fig. 12.2, for which the output voltage range is ±5 V, is required using $V_{CC} = 10$ V. The circuit is to be designed such that the current variation in the emitter-follower transistor is no greater than a factor of 15, for load resistances as low as 100 Ω. What is the value of R required? Find the incremental voltage gain of the resulting follower at $v_O = +5$, 0, and –5 V, with a 100-Ω load. What is the percentage change in gain over this range of v_O?

***12.5** Consider the operation of the follower circuit of Fig. 12.2 for which $R_L = V_{CC}/I$, when driven by a square wave such that the output ranges from $+V_{CC}$ to $-V_{CC}$ (ignoring $V_{CE\,sat}$). For this situation, sketch the equivalent of Fig. 12.4 for v_O, i_{C1}, and p_{D1}. Repeat for a square-wave output that has peak levels of $\pm V_{CC}/2$. What is the average power dissipation in Q_1 in each case? Compare these results to those for sine waves of peak amplitude V_{CC} and $V_{CC}/2$, respectively.

12.6 Consider the situation described in Problem 12.5. For square-wave outputs having $\pm V_{CC}$ levels and $\pm\frac{1}{2}V_{CC}$ levels, and for sine waves of the same peak-to-peak values, find the average power loss in the current-source transistor Q_2.

12.7 Reconsider the situation described in Exercise 12.3 for variation in V_{CC}—specifically for $V_{CC} = 16$ V, 12 V, 10 V, and 8 V. Assume $V_{CE\,sat}$ is nearly zero. What is the power-conversion efficiency in each case?

D 12.8 The emitter-follower output stage of Fig. 12.2 is designed to provide a maximum output swing of $\pm\hat{V}$ volts, across the load R_L. Neglecting the saturation voltage, what are the minimum required values of V_{CC} and I? Now, if the output voltage is a sine wave of peak amplitude $(\hat{V}/2)$, what is the power-conversion efficiency realized?

Section 12.3: Class B Output Stage

12.9 Consider the circuit of a complementary-BJT class B output stage. For what amplitude of input signal does the crossover distortion represent a 10% loss in peak amplitude?

12.10 Consider the feedback configuration with a class B output stage shown in Fig. 12.9. Let the amplifier gain $A_0 = 100$ V/V. Derive an expression for v_O versus v_I, assuming that $|V_{BE}| = 0.7$ V. Sketch the transfer characteristic v_O versus v_I, and compare it with that without feedback.

SIM **12.11** Consider the class B output stage, using MOSFETs, shown in Fig. P12.11. Let the devices have $|V_t| = 0.5$ V and $\mu C_{ox} W/L = 2$ mA/V^2. With a 10-kHz sine-wave input of 5-V peak and a high value of load resistance, what peak output would you expect? What fraction of the sine-wave period does the crossover interval represent? For what value of load resistor is the peak output voltage reduced to half the input?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

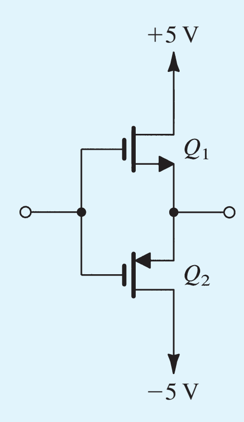


Figure P12.11

12.12 Consider the complementary-BJT class B output stage and neglect the effects of finite V_{BE} and $V_{CE\text{sat}}$. For ±10-V power supplies and an 8-Ω load resistance, what is the maximum sine-wave output power available? What supply power corresponds? What is the power-conversion efficiency? For output signals of half this amplitude, find the output power, the supply power, and the power-conversion efficiency.

D 12.13 A class B output stage operates from ±10-V supplies. Assuming relatively ideal transistors, what is the output voltage for maximum power-conversion efficiency? What is the output voltage for maximum device dissipation? If each of the output devices is individually rated for 2-W dissipation, and a factor-of-2 safety margin is to be used, what is the smallest value of load resistance that can be tolerated, if operation is always at full output voltage? If operation is allowed at half the full output voltage, what is the smallest load permitted? What is the greatest possible output power available in each case?

D 12.14 A class B output stage is required to deliver an average power of 50 W into an 8-Ω load. The power supply should be 4 V greater than the corresponding peak sine-wave output voltage. Determine the power-supply voltage required (to the nearest volt in the appropriate direction), the peak current from each supply, the total supply power, and the power-conversion efficiency. Also, determine the maximum possible power dissipation in each transistor for a sine-wave input.

12.15 Consider the class B BJT output stage with a square-wave output voltage of amplitude $\hat{V}_o$ across a load R_L and employing power supplies $\pm V_{SS}$. Neglecting the effects of finite V_{BE} and $V_{CE\text{sat}}$, determine the load power, the supply power, the power-conversion efficiency, the maximum attainable power-conversion efficiency and the corresponding value of $\hat{V}_o$, and the maximum available load power. Also find the value of $\hat{V}_o$ at which the power dissipation in the transistors reaches its peak, and the corresponding value of power-conversion efficiency.

12.16 Sketch a graph for the small-signal voltage gain of the class B circuit of Fig. 12.5 as a function of v_I, for v_I both positive and negative.

Section 12.4: Class AB Output Stage

12.17 A class AB output stage, such as that in Fig. 12.11, utilizing transistors with $I_S = 10^{-14}$ A, is biased at a quiescent current $I_Q = 1$ mA. Find V_{BB}, the output resistance R_{out} at $v_I = 0$, and the corresponding small-signal voltage gain. The load resistance $R_L = 100\ \Omega$. What does the incremental gain become when $v_O = 10$ V?

D 12.18 Design the quiescent current of a class AB BJT output stage so that the incremental voltage gain for v_I in the vicinity of the origin is in excess of 0.97 V/V for loads larger than 100 Ω. Assume that the BJTs have V_{BE} of 0.7 V at a current of 100 mA and determine the value of V_{BB} required.

D 12.19 A class AB output stage, such as that in Fig. 12.11, drives a load resistance R_L of 100 Ω. What bias current I_Q will serve to limit the variation in the small-signal voltage gain to 5% as i_L changes from 0 to 50 mA?

12.20 For the class AB output stage considered in Example 12.3, add two columns to the table of results as follows: the total input current drawn from v_I (i_I, mA); and the large-signal input resistance $R_{\text{in}} \equiv v_I/i_I$. Assume $\beta_N = \beta_P = \beta = 49$. Compare the values of R_{in} to the approximate value obtained using the resistance-reflection rule, $R_{\text{in}} \simeq \beta R_L$.

12.21 In this problem we investigate an important trade-off in the design of the class AB output stage of Fig. 12.11: Increasing the quiescent current I_Q reduces the nonlinearity of the transfer characteristic at the expense of increased quiescent power dissipation. As a measure of nonlinearity, we use the maximum deviation of the stage incremental gain, which occurs at $v_O = 0$, namely,

$$\epsilon = 1 - v_o/v_i\big|_{v_O=0}$$

(a) Show that ϵ is given by

$$\epsilon = \frac{V_T/2I_Q}{R_L + (V_T/2I_Q)}$$

which for $2I_QR_L \gg V_T$ can be approximated by

$$\epsilon \simeq V_T/2I_QR_L$$

(b) If the stage is operated from power supplies of $\pm V_{CC}$, find the quiescent power dissipation, P_D.

(c) Show that for given V_{CC} and R_L, the product of the quiescent power dissipation and the gain error is a constant given by

$$\epsilon P_D \simeq V_T\left(\frac{V_{CC}}{R_L}\right)$$

(d) For $V_{CC} = 10$ V and $R_L = 100\ \Omega$, find the required values of P_D and I_Q if ϵ is to be 5%, 2%, and 1%.

***12.22** A class AB output stage, resembling that in Fig. 12.11 but utilizing a single supply of +10 V and biased at $V_I = 6$ V, is capacitively coupled to a 100-Ω load. For transistors for which $|V_{BE}| = 0.7$ V at 1 mA and for a bias voltage $V_{BB} = 1.4$ V, what quiescent current results? For a step change in output from 0 to –1 V, what input step is required? Assuming transistor-saturation voltages of zero, find the largest possible positive-going and negative-going steps at the output.

Section 12.5: Biasing the Class AB Circuit

D 12.23 Consider the diode-biased class AB circuit of Fig. 12.14. For $I_{BIAS} = 200\ \mu$A, find the relative size (n) that should be used for the output devices (in comparison to the biasing devices) to ensure that an output resistance of 8 Ω or less is obtained in the quiescent state. Neglect the resistance of the biasing diodes.

D*12.24 A class AB output stage using a two-diode bias network as shown in Fig. 12.14 utilizes diodes having the same junction area as the output transistors. For $V_{CC} = 10$ V, $I_{BIAS} = 1$ mA, $R_L = 100\ \Omega$, $\beta_N = 50$, and $|V_{CEsat}| = 0$ V, what is the quiescent current? What are the largest possible positive and negative output signal levels? To achieve a positive peak output level equal to the negative peak level, what value of β_N is needed if I_{BIAS} is not changed? What value of I_{BIAS} is needed if β_N is held at 50? For this value, what does I_Q become?

D 12.25 It is required to evaluate the small-signal input resistance and small-signal voltage gain of the class AB output stage in Fig. 12.14. To simplify matters, assume the small-signal resistances of D_1 and D_2 to be negligibly small. Replace each of Q_N and Q_P with its hybrid-π model and neglect r_o. Hence show that the class AB stage is equivalent, from a small-signal point of view, to an emitter-follower transistor whose $r_\pi = r_{\pi N} \| r_{\pi P}$ and $g_m = g_{mN} + g_{mP}$, and hence $r_e = r_{eN} \| r_{eP}$ and $\beta = (g_{mN} + g_{mP})(r_{\pi N} \| r_{\pi P})$. Now show that

$$\frac{v_o}{v_i} = \frac{R_L}{R_L + (r_{eN} \| r_{eP})}$$

and

$$R_{in} \simeq \beta[R_L + (r_{eN} \| r_{eP})]$$

12.26 Figure P12.26 shows a class AB output stage with a common-emitter transistor added to increase the voltage gain and reduce the current that v_I has to supply. Neglecting the small-signal resistances of D_1 and D_2, find the small-signal voltage gain v_o/v_i. (*Hint*: Use the expressions for voltage gain and input resistance of the class AB stage without Q_3, given in the statement for Problem 12.25.)

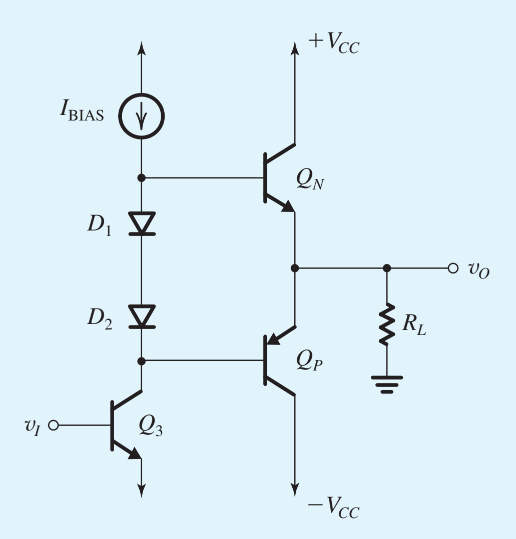


Figure P12.26

12.27 It is required to find an expression for the output resistance R_{out} of the class AB output stage in Fig. P12.26. Toward that end, neglect the small-signal resistance of each of D_1 and D_2 and assume the current source supplying I_{BIAS} has an

output resistance R_{BIAS}. Transistors Q_N and Q_P are equivalent to a single transistor with $r_\pi = r_{\pi N} \| r_{\pi P}$, $r_e = r_{eN} \| r_{eP}$, and $g_m = g_{mN} + g_{mP}$.

****12.28** A class AB output stage using a two-diode bias network as shown in Fig. 12.14 utilizes diodes having the same junction area as the output transistors. At a room temperature of about 20°C the quiescent current is 1 mA and $|V_{BE}| = 0.6$ V. Through a manufacturing error, the thermal coupling between the output transistors and the biasing diode-connected transistors is omitted. After some output activity, the output devices heat up to 70°C while the biasing devices remain at 20°C. Thus, while the V_{BE} of each device remains unchanged, the quiescent current in the output devices increases. To calculate the new current value, recall that there are two effects: I_S increases by about 14%/°C and $V_T = kT/q$ changes, where $T = 273° +$ temperature in °C, and $V_T = 25$ mV only at 20°C. However, you may assume that β_N remains almost constant. This assumption is based on the fact that β increases with temperature but decreases with current. What is the new value of I_Q? If the power supply is ±20 V, what additional power is dissipated? If thermal runaway occurs, and the temperature of the output transistors increases by 10°C for every watt of additional power dissipation, what additional temperature rise and current increase result?

D 12.29 Repeat Example 12.5 for the situation in which the peak positive output current is 250 mA. Use the same general approach to safety margins. What are the values of R_1 and R_2 you have chosen?

****12.30** A V_{BE} multiplier is designed with equal resistances for nominal operation at a terminal current of 1 mA, with half the current flowing in the bias network. The initial design is based on $\beta = \infty$ and $V_{BE} = 0.7$ V at 1 mA.

(a) Find the required resistor values and the terminal voltage.
(b) Find the terminal voltage that results when the terminal current increases to 2 mA. Assume $\beta = \infty$.
(c) Repeat (b) for the case the terminal current becomes 10 mA.
(d) Repeat (c) using the more realistic value of $\beta = 100$.

***12.31** By replacing the transistor in the V_{BE} multiplier by its hybrid-π, small-signal model (with r_o neglected), show that the incremental resistance between the two terminals of the multiplier is given by

$$r = \frac{R_2 + (R_1 \| r_\pi)}{1 + g_m (R_1 \| r_\pi)}$$

Evaluate r for the case $R_1 = R_2 = 1.2$ kΩ, with the transistor operating at $I_C = 1$ mA and having $\beta = 100$.

Section 12.6: Variations on the Class AB Configuration

12.32 Use the results given in the answer to Exercise 12.9 to determine the input current of the circuit in Fig. 12.17 for $v_I = 0$ and ±10 V with infinite and 100-Ω loads.

12.33 For the circuit in Fig 12.17, operated near $v_I = 0$ and fed with a signal source having zero resistance, show that the output resistance is given by

$$R_{out} = \frac{1}{2}\left[R_3 + r_{e3} + \left(R_1 \| r_{e1}\right)/\left(\beta_3 + 1\right)\right]$$

Assume that the top and bottom halves of the circuit are perfectly matched.

D ***12.34 Consider the circuit of Fig. 12.17 in which Q_1 and Q_2 are matched, and Q_3 and Q_4 are matched but have three times the junction area of the others. Resistors R_3 and R_4 also are matched. For $V_{CC} = 10$ V, find values for resistors R_1 through R_4 that allow for a base current of at least 10 mA in Q_3 (and Q_4) at $v_I = +5$ V ($v_I = -5$ V), when a load demands it, with at most a 2-to-1 variation in currents in Q_1 (and Q_2). The quiescent current in Q_3 is to be 40 mA. Let $\beta_{1,2} \geq 150$ and $\beta_{3,4} \geq 50$. For input voltages around 0 V, estimate the output resistance of the overall follower driven by a source having zero resistance. For an input voltage of +1 V and a load resistance of 2 Ω, what output voltage results? Q_1 and Q_2 have $|V_{BE}|$ of 0.7 V at a current of 10 mA.

12.35 Figure P12.35 shows a variant of the class AB circuit of Fig. 12.17. Assume that all four transistors are matched and have $\beta = 100$.

(a) For $v_I = 0$, find the quiescent current in Q_3 and Q_4, the input current i_I, and the output voltage v_O.
(b) Since the circuit has perfect symmetry, the small-signal performance around $v_I = 0$ can be determined by considering either the top or bottom half of the circuit only. In this case, the load on the half-circuit must be $2R_L$, the input resistance found is $2R_{in}$, and the output resistance

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

found is $2R_{out}$. Using this approach, find R_{in}, v_o/v_i, and R_{out} (assuming that the circuit is fed with a zero-resistance source).

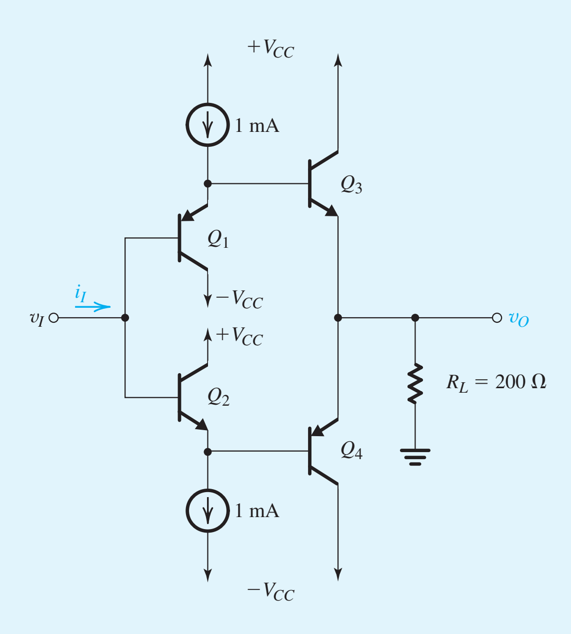


Figure P12.35

12.36 For the Darlington configuration shown in Fig. 12.18, show that for $\beta_1 \gg 1$ and $\beta_2 \gg 1$:

(a) The equivalent composite transistor has $\beta \simeq \beta_1\beta_2$.
(b) If the composite transistor is operated at a current I_C, then Q_2 will be operating at a collector current approximately equal to I_C, and Q_1 will be operating at a collector current approximately equal to I_C/β_2.
(c) The composite transistor has a base–emitter voltage $V_{BE} \simeq 2V_T \ln(I_C/I_S) - V_T \ln(\beta_2)$, where I_S is the saturation current of each of Q_1 and Q_2.
(d) The composite transistor has an equivalent $r_\pi \simeq 2\beta_1\beta_2(V_T/I_C)$.
(e) The composite transistor has an equivalent $g_m \simeq \frac{1}{2}(I_C/V_T)$.

***12.37** For the circuit in Fig. P12.37 in which the transistors have $V_{BE} = 0.7$ V and $\beta = 100$:

(a) Find the dc collector current for each of Q_1 and Q_2.
(b) Find the small-signal current i_c that results from an input signal v_i, and hence find the voltage gain v_o/v_i.
(c) Find the input resistance R_{in}.

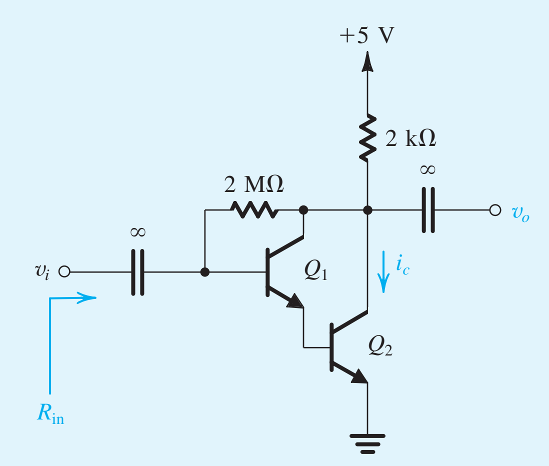


Figure P12.37

SIM ****12.38** The BJTs in the circuit of Fig. P12.38 have $\beta_P = 10$, $\beta_N = 100$, $|V_{BE}| = 0.7$ V, and $|V_A| = 100$ V.

(a) Find the dc collector current of each transistor and the value of V_C.
(b) Replacing each BJT with its hybrid-π model, show that

$$\frac{v_o}{v_i} \simeq g_{m1}\left[r_{o1} \,\|\, \beta_N(r_{o2} \,\|\, R_f)\right]$$

(c) Find the values of v_o/v_i and R_{in}.

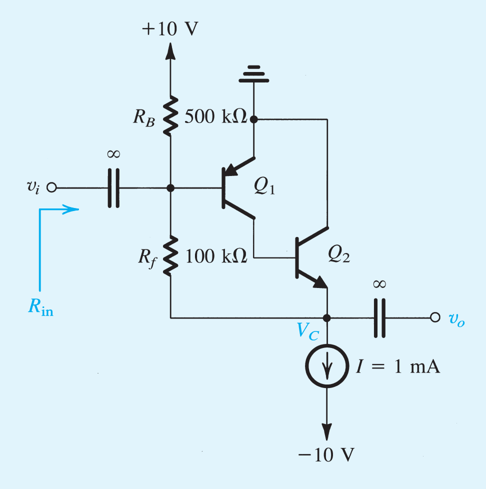


Figure P12.38

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

D **12.39 Consider the compound-transistor class AB output stage shown in Fig. 12.20 in which Q_2 and Q_4 are matched transistors with $V_{BE} = 0.7$ V at 10 mA and $\beta = 100$, Q_1 and Q_5 have $V_{BE} = 0.7$ V at 1-mA currents and $\beta = 100$, and Q_3 has $V_{EB} = 0.7$ V at a 1-mA current and $\beta = 10$. Design the circuit for a quiescent current of 2 mA in Q_2 and Q_4, I_{BIAS} that is 100 times the standby base current in Q_1, and a current in Q_5 that is nine times that in the associated resistors. Find the values of the input voltage required to produce outputs of ±10 V for a 1-kΩ load. Use V_{CC} of 15 V.

***12.40** Figure P12.40 shows a variant on the class AB amplifier known as class G. Here, in addition to the

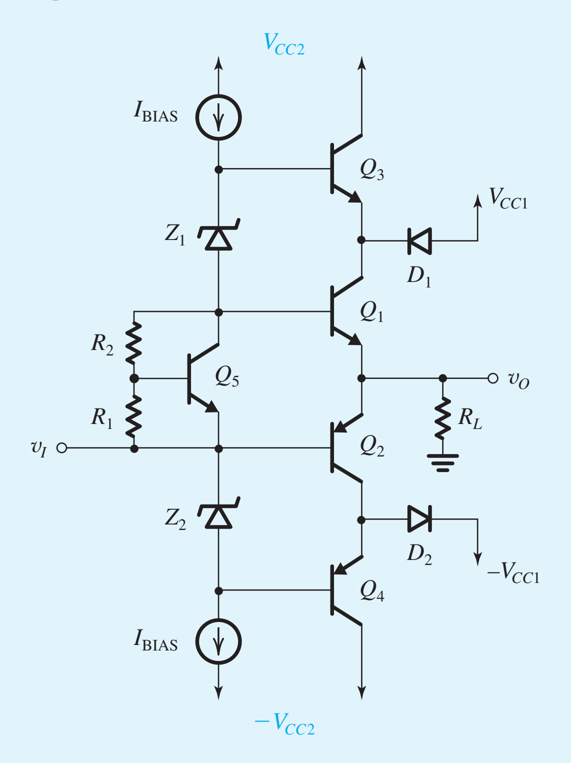


Figure P12.40

normal power supply $\pm V_{CC1}$, the circuit is equipped with a higher voltage supply $\pm V_{CC2}$. The latter supply is utilized only infrequently. The circuit operates as follows. Normally, D_1 and D_2 are turned on and thus connect the $\pm V_{CC1}$ supply to the class AB stage transistors Q_1 and Q_2. Simultaneously, Q_3 and Q_4 are off. For v_I positive and exceeding a certain threshold, Q_3 turns on, D_1 turns off, and Q_1 is then effectively operating from the higher voltage supply V_{CC2}. This continues as long as v_I is larger than the specified threshold. As v_I decreases below the threshold value, Q_3 is turned off and D_1 turns on, thus connecting Q_1 to its normal supply V_{CC1}. A similar process happens in the negative direction, with D_2 and Q_4 taking the place of D_1 and Q_3. Let $V_{CC1} = 35$ V, $V_{CC2} = 70$ V, $V_{Z1} = 3.3$ V, and the voltage of the V_{BE} multiplier $V_{BB} = 1.2$ V.

(a) Find the positive threshold value of v_I at which Q_3 is turned on.
(b) If for 95% of the time v_I is in the vicinity of 30 V and only 5% of the time it is in the vicinity of 65 V, use Eq. (12.19) to estimate the average power dissipated in the transistors, P_D. Compare to the value of P_D dissipated in a class AB stage operated from a ±70 V supply.

12.41 Repeat Exercise 12.11 for a design variation in which transistor Q_5 is increased in size by a factor of 20, all other conditions remaining the same.

12.42 Repeat Exercise 12.11 for a design in which the limiting output current and normal peak current are 100 mA and 75 mA, respectively.

D 12.43 The circuit shown in Fig. P12.43 operates in a manner analogous to that in Fig. 12.21 to limit the output

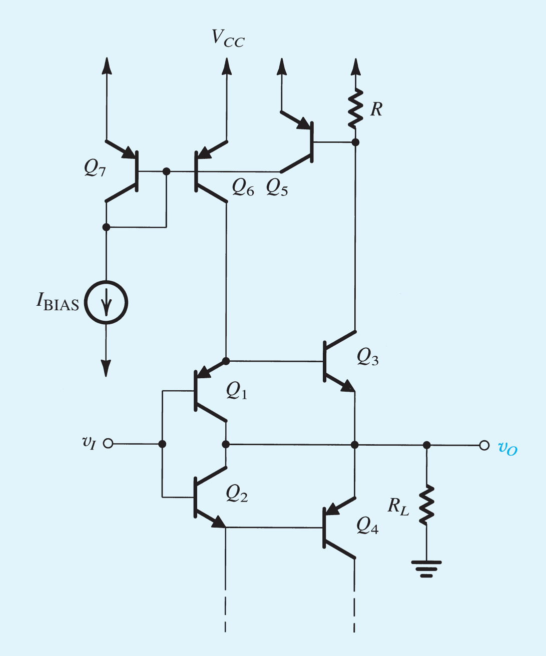


Figure P12.43

current from Q_3 in the event of a short circuit or other mishap. It has the advantage that the current-sensing resistor R does

not appear directly at the output. Find the value of R that causes Q_5 to turn on and absorb all of $I_{\text{BIAS}} = 2$ mA, when the current being sourced reaches 100 mA. For Q_5, $I_S = 10^{-14}$ A. If the normal peak output current is 75 mA, find the voltage drop across R and the collector current in Q_5.

D 12.44 Consider the thermal shutdown circuit shown in Fig. 12.22. At 25°C, Z_1 is a 6.8-V zener diode with a TC of 2 mV/°C, and Q_1 and Q_2 are BJTs that display V_{BE} of 0.7 V at a current of 100 μA and have a TC of –2 mV/°C. Design the circuit so that at 125°C, a current of 200 μA flows in each of Q_1 and Q_2. What is the current in Q_2 at 25°C?

Section 12.7: CMOS Class AB Output Stages

D 12.45 (a) Show that for the class AB circuit in Fig. 12.23, the small-signal output resistance in the quiescent state is given by

$$R_{\text{out}} \simeq \frac{1}{g_{mn} + g_{mp}}$$

which for matched devices becomes

$$R_{\text{out}} = \frac{1}{2g_m}$$

(b) For a circuit that utilizes MOSFETs with $|V_t| = 0.5$ V and $k'(W/L) = 200$ mA/V^2, find the voltage V_{GG} that results in $R_{\text{out}} = 20\ \Omega$. Also, find I_Q.

D 12.46 (a) For the circuit in Fig. 12.23 in which Q_1 and Q_2 are matched, Q_N and Q_P are matched, and all devices have the same $|V_t|$, show that the small-signal voltage gain at the quiescent condition is given by

$$\frac{v_o}{v_i} = \frac{R_L}{R_L + (2/g_m)}$$

where g_m is the transconductance of each of Q_N and Q_P and where channel-length modulation is neglected.
(b) For the case $I_{\text{BIAS}} = 0.2$ mA, $R_L = 1$ kΩ, $k_n = k_p = nk_1 = nk_2$, where $k = \mu C_{ox}(W/L)$, and $k_1 = 20$ mA/V^2, find the ratio n that results in an incremental gain of 0.98. Also find the quiescent current I_Q.

D 12.47 Design the circuit of Fig. 12.23 to operate at $I_Q = 1$ mA with $I_{\text{BIAS}} = 0.1$ mA. Let $\mu_n C_{ox} = 250\ \mu$A/V^2, $\mu_p C_{ox} = 100\ \mu$A/V^2, $V_{tn} = -V_{tp} = 0.45$ V, and $V_{DD} = V_{SS} = 2.5$ V. Design so that Q_1 and Q_2 are matched and Q_N and Q_P are matched, and that in the quiescent state each operates at an overdrive voltage of 0.15 V.

(a) Specify the W/L ratio for each of the four transistors.
(b) In the quiescent state with $v_O = 0$, what must v_I be?
(c) If Q_N is required to supply a maximum load current of 10 mA, find the maximum allowable output voltage. Assume that the transistor supplying I_{BIAS} needs a minimum of 0.2 V to operate properly.

D 12.48 Consider the design of the class AB output stage of Fig. 12.23 for the following conditions. The stage is operated from ±2.5-V power supplies and is required to provide a minimum output voltage swing of ±1.5 V while supplying a maximum current equal to 10 times the quiescent current I_Q. Assume that Q_N and Q_P are matched and Q_1 and Q_2 are matched, that all devices have $|V_t| = 0.5$ V, and that in the quiescent state all transistors are operated at the same overdrive voltage. What is the value of V_{OV} required, and what V_{GG} is needed?

12.49 The class AB output stage in Fig. 12.24 utilizes two matched transistors with $k_n = k_p = 200$ mA/V^2 and is operated from ±2.5-V power supplies. If the stage is required to supply a maximum current of ±20 mA, what is the output voltage swing realized?

12.50 For the CMOS output stage of Fig. 12.25 with $I_Q = 2$ mA, $|V_{OV}| = 0.2$ V for each of Q_P and Q_N at the quiescent point, and $\mu = 5$, find the output resistance at the quiescent point.

12.51 (a) Show that for the CMOS output stage of Fig. 12.25,

$$|\text{Gain error}| = \frac{R_{\text{out}}}{R_L}$$

(b) For a stage that drives a load resistance of 100 Ω with a gain error of less than 3%, find the overdrive voltage at which Q_P and Q_N should be operated. Let $I_Q = 2.5$ mA and $\mu = 5$.

12.52 Show that in the CMOS class AB common-source output stage (Fig. 12.25), Q_N turns off when $v_O = 4I_Q R_L$ and Q_P turns off when $v_O = -4I_Q R_L$. This is equivalent to saying that one of the transistors turns off when $|i_L|$ reaches $4I_Q$.

D *12.53 It is required to design the circuit of Fig. 12.25 to drive a load resistance of 50 Ω while exhibiting an output resistance, around the quiescent point, of 2.5 Ω.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

Operate Q_N and Q_P at $I_Q = 1.5$ mA and $|V_{OV}| = 0.15$ V. The technology utilized is specified to have $k_n' = 250\ \mu A/V^2$, $k_p' = 100\ \mu A/V^2$, $V_{tn} = -V_{tp} = 0.5$ V, and $V_{DD} = V_{SS} = 2.5$ V.

(a) Specify (W/L) for each of Q_N and Q_P.
(b) Specify the required value of μ.
(c) What is the expected error in the stage gain?
(d) In the quiescent state, what dc voltage must appear at the output of each of the error amplifiers?
(e) At what value of positive v_O will Q_P be supplying all the load current? Repeat for negative v_O and Q_N supplying all the load current.
(f) What is the linear range of v_O?

***12.54** Figure P12.54 shows a class AB output stage utilizing a pair of complementary MOSFETs (Q_N, Q_P) and employing BJTs for biasing and in the driver stage. The latter consists of complementary Darlington emitter followers formed by Q_1 through Q_4 and has the low output resistance necessary for driving the output MOSFETs at high speeds. Of special interest is the bias circuit utilizing two V_{BE} multipliers formed by Q_5 and Q_6 and their associated resistors. Transistor Q_6 is placed in direct thermal contact with the output transistors and thus has the same temperature as that of Q_N and Q_P.

(a) Show that V_{GG} is given by

$$V_{GG} = \left(1 + \frac{R_3}{R_4}\right) V_{BE6} + \left(1 + \frac{R_1}{R_2}\right) V_{BE5} - 4V_{BE}$$

(b) Noting that V_{BE6} is thermally coupled to the output devices while the other BJTs remain at constant temperature, show that

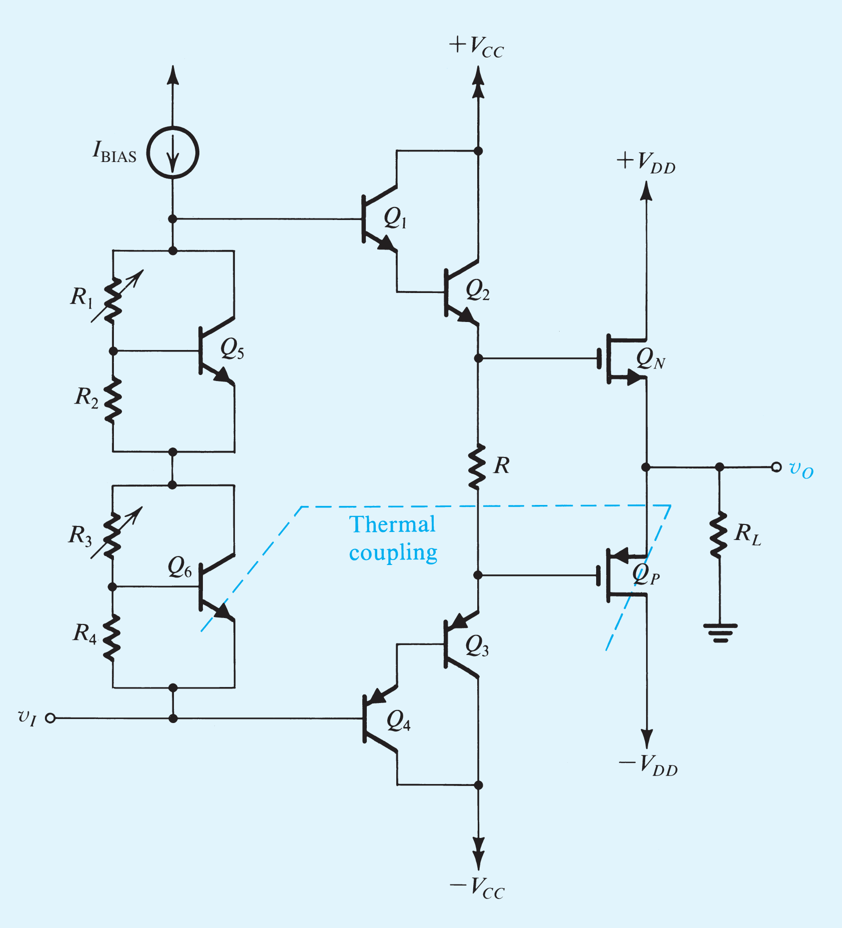


Figure P12.54

$$\frac{\partial V_{GG}}{\partial T} = \left(1 + \frac{R_3}{R_4}\right)\frac{\partial V_{BE6}}{\partial T}$$

(c) To keep the overdrive voltages of Q_N and Q_P, and hence their quiescent current, constant with temperature variation, $\partial V_{GG}/\partial T$ is made equal to $\partial(V_{tN} + V_{tP})/\partial T$. Find R_3/R_4 that provides this temperature stabilization when $|V_t|$ changes by -3 mV/°C and $\partial V_{BE}/\partial T = -2$ mV/°C.

(d) Using the value of R_3/R_4 found in (c) and assuming that the nominal value of V_{BE} is 0.7 V and that the MOSFETs have $|V_t| = 3$ V and $\mu C_{ox}(W/L) = 2$ A/V^2, find $|V_{GS}|$, V_{GG}, R, and R_1/R_2 to establish a quiescent current of 100 mA in the output transistors and 20 mA in the driver stage.

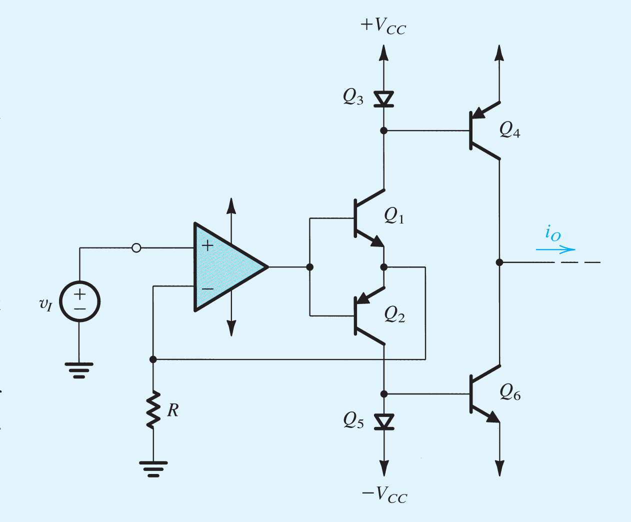


Figure P12.58

Section 12.8: IC Power Amplifiers

D 12.55 In the power-amplifier circuit of Fig. 12.29, two resistors are important in controlling the overall voltage gain. Which are they? Which controls the gain alone? Which affects both the dc output level and the gain? A new design is being considered in which the output dc level is approximately $\frac{2}{3}V_S$ (rather than approximately $\frac{1}{2}V_S$) with a gain of 50 (as before). What changes are needed?

12.56 Consider the front end of the circuit in Fig. 12.29. For $V_S = 22$ V, calculate approximate values for the bias currents in Q_1 through Q_6. Assume $\beta_{npn} = 100$, $\beta_{pnp} = 20$, and $|V_{BE}| = 0.7$ V. Also find the dc voltage at the output.

D 12.57 It is required to use the LM380 power amplifier to drive an 8-Ω loudspeaker while limiting the maximum possible device dissipation to 2 W. Use the graph of Fig. 12.31 to determine the maximum possible power-supply voltage that can be used. (Use only the given graphs; do not interpolate.) If the maximum allowed THD is to be 3%, what is the maximum possible load power? To deliver this power to the load what peak-to-peak output sinusoidal voltage is required?

12.58 For the circuit in Fig. P12.58, assuming all transistors to have large β, show that $i_O = v_I/R$. [This voltage-to-current converter is an application of a versatile circuit building block known as the **current conveyor**; see Sedra and Roberts (1990).] For $\beta = 100$, by what approximate percentage is i_O actually lower than this ideal value?

D 12.59 For the bridge amplifier of Fig. 12.32, let $R_1 = R_3 = 10$ kΩ. Find R_2 and R_4 to obtain an overall gain of 8 V/V.

D 12.60 An alternative bridge amplifier configuration, with high input resistance, is shown in Fig. P12.60. [Note the similarity of this circuit to the front end of the instrumentation amplifier circuit shown in Fig. 2.20(b).] What is the gain v_O/v_I? For op amps (using ±15-V supplies) that limit at ±13 V, what is the largest sine wave you can provide across R_L? Using 1 kΩ as the smallest resistor, find resistor values that make $v_O/v_I = 8$ V/V. Make sure that the signals at the outputs of the two amplifiers are complementary.

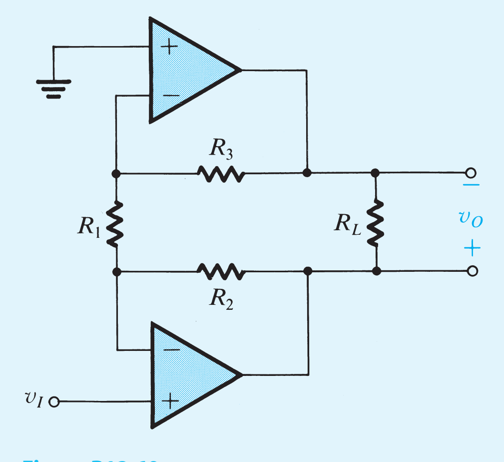


Figure P12.60

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

Section 12.9 Class D Power Amplifiers

12.61 Sketch diagrams resembling those in Figs. 12.33(a), (b). Let v_T have ± 10 V peaks and assume v_A is a sine wave with 5-V peak amplitude. Let the frequency of v_T be 5 times that of v_A. The comparator output levels are ± 10 V.

12.62 A pulse waveform swinging between ± 10 V has a duty ratio of 0.65. What is its average value? If the duty ratio is changed to 0.35, what does the average value become?

12.63 For the circuit in Fig. 12.34(b):

(a) If v_A is a sine wave, what is the maximum power supplied to a load of resistance R, in terms of V_{DD}?

(b) The power loss is mostly due to the repeated charging and discharging of a capacitance C across the load. It can be shown that this switching power is given by $4f_s C V_{DD}^2$. Find an expression for the power-conversion efficiency η and evaluate the value of η for the case $f_s = 250$ kHz and $C = 1000$ pF.

Section 12.10 Power Transistors

12.64 A power MOSFET is specified to have $I_{D\max} = 5$ A, $V_{DS\max} = 50$ V, and $P_{D\max} = 50$ W.

(a) Sketch the SOA boundaries.

(b) If the MOSFET is used in the common-source configuration as shown in Fig. P12.64, show that the maximum current occurs when $V_{DS} = 0$, the maximum V_{DS} occurs when $I_D = 0$, and the maximum power dissipation occurs when $V_{DS} = V_{DD}/2$.

(c) For $V_{DD} = 40$ V, find the smallest resistance R for which the operating point is always within the SOA. What are the corresponding values of $I_{D\max}$ and $P_{D\max}$?

(d) Repeat (c) for $V_{DD} = 30$ V.

(e) Repeat (c) for $V_{DD} = 15$ V.

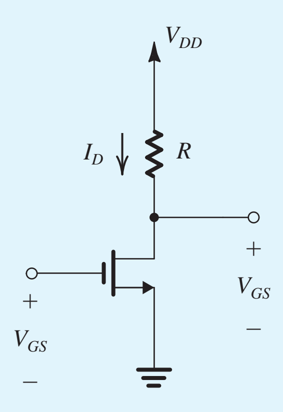


Figure P12.64

D 12.65 A particular transistor having a thermal resistance $\theta_{JA} = 2.5$°C/W is operating at an ambient temperature of 30°C with a collector–emitter voltage of 20 V. If long life requires a maximum junction temperature of 130°C, what is the corresponding device power rating? What is the greatest average collector current that should be considered?

12.66 A particular transistor has a power rating at 25°C of 10 W, and a maximum junction temperature of 150°C. What is its thermal resistance? What is its power rating when operated at an ambient temperature of 50°C? What is its junction temperature when dissipating 5 W at an ambient temperature of 50°C?

12.67 A power transistor operating at an ambient temperature of 50°C, and an average emitter current of 3 A, dissipates 20 W. If the thermal resistance of the transistor is known to be less than 3°C/W, what is the highest junction temperature you would expect? If the transistor V_{BE} measured using a pulsed emitter current of 3 A at a junction temperature of 25°C is 0.80 V, what average V_{BE} would you expect under normal operating conditions? (Use a temperature coefficient of –2 mV/°C.)

12.68 For a particular application of the transistor specified in Example 12.7, extreme reliability is essential. To improve reliability, the maximum junction temperature is to be limited to 100°C. What are the consequences of this decision for the conditions specified?

12.69 A power transistor is specified to have a maximum junction temperature of 150°C. When the device is operated at this junction temperature with a heat sink, the case temperature is found to be 97°C. The case is attached to the heat sink with a bond having a thermal resistance $\theta_{CS} = 0.5$°C/W and the thermal resistance of the heat sink $\theta_{SA} = 0.1$°C/W. If the ambient temperature is 25°C, what is the power being dissipated in the device? What is the thermal resistance of the device, θ_{JC}, from junction to case?

12.70 A power transistor for which $T_{J\max} = 180$°C can dissipate 50 W at a case temperature of 30°C. If it is connected to a heat sink using an insulating washer for which the thermal resistance is 0.6°C/W, what heat-sink temperature is necessary to ensure safe operation at 30 W? For an ambient temperature of 27°C, what heat-sink thermal resistance is required? If, for a particular extruded-aluminum-finned heat sink, the thermal resistance in still air is 6°C/W per centimeter of length, how long a heat sink is needed?

CHAPTER 13

Operational-Amplifier Circuits

IN THIS CHAPTER YOU WILL LEARN

1. The design and analysis of the two basic CMOS op-amp architectures: the two-stage circuit and the single-stage, folded-cascode circuit.
2. The complete circuit of an analog IC classic: the 741 BJT op amp. Though over 40 years old, the 741 circuit includes so many interesting and useful design techniques that its study is still a must.
3. Interesting and useful applications of negative feedback within op-amp circuits to achieve bias stability and increased CMRR.
4. How to break a large analog circuit into its recognizable blocks to be able to make the analysis amenable to a pencil-and-paper approach, which is the best way to learn design.
5. Some of the modern techniques employed in the design of low-voltage, single-supply BJT op amps.
6. Most important, how the different topics we studied in the preceding chapters come together in the design of the most important analog IC, the op amp.

Introduction

In this chapter, we shall study the internal circuitry of the most important analog IC, namely, the operational amplifier. The terminal characteristics and some circuit applications of op amps were covered in Chapter 2. Here, our objective is to expose the reader to some of the ingenious techniques that have evolved over the years for combining elementary analog circuit building blocks to realize a complete op amp. We shall study both CMOS and bipolar op amps.

The CMOS op-amp circuits considered find application primarily in the design of analog and mixed-signal VLSI circuits. Because these op amps are usually designed with a specific application in mind, they can be optimized to meet a subset of the list of desired specifications, such as high dc gain, wide bandwidth, or large output-signal swing. For instance, many CMOS op amps are utilized within an IC and do not connect to the outside terminals of the chip. As a result, the loads on their outputs are usually limited to small capacitances of at most a few picofarads. Internal CMOS op amps therefore do not need to have low output resistances, and their design rarely incorporates an output stage. Also, if the op-amp input terminals are not connected to the chip terminals, there will be no danger of static charge damaging the gate oxide of the input MOSFETs. Hence, internal CMOS op amps do not need input

clamping diodes for gate protection and thus do not suffer from the leakage effects of such diodes. In other words, the advantage of near-infinite input resistance of the MOSFET is fully realized.

While CMOS op amps are extensively used in the design of VLSI systems, the BJT remains the device of choice in the design of general-purpose op amps. These are op amps that are utilized in a wide variety of applications and are designed to fit a wide range of specifications. As a result, the circuit of a general-purpose op amp represents a compromise among many performance parameters. We shall study in detail one such circuit, the 741-type op amp. Although the 741 has been available for over 40 years, its internal circuit remains as relevant and interesting today as it ever was. Nevertheless, changes in technology have introduced new requirements, such as the need for general-purpose op amps that operate from a single power supply of only 2 V to 3 V. These new requirements have given rise to exciting challenges to op-amp designers. The result has been a wealth of new ideas and design techniques. We shall present a sample of these modern design techniques in the last section.

In addition to exposing the reader to some of the ideas that make analog IC design such an exciting topic, this chapter should serve to tie together many of the concepts and methods studied thus far.

THE GENIE OF ANALOG:

The need for precision in analog ICs supported a generation of highly skilled circuit and process engineers. Their creative approaches to overcoming the limitations of available technology led to celebrity status among their peers. Perhaps most famous of all was Robert Widlar, who teamed with process engineer Dave Talbert to bootstrap the analog business in the 1960s with highly successful designs for Fairchild and National Semiconductor. Widlar's pranks, including threatening to cut through bureaucracy with an axe, and buying a sheep to trim National's unkempt lawns, remain Silicon Valley legends.

Robert John Widlar was already a legendary chip designer at age 33 (but died at 53), and a pioneer of linear analog integrated-circuit design: the creator of the Widlar current source, the Widlar bandgap voltage reference, the Widlar output stage, and a host of op-amp designs, including the first mass-produced operational amplifier ICs (Fairchild μA702, μA709), the first integrated voltage regulator (μA723, National LM100), the first fully internally compensated operational amplifier (LM101), the field-effect input (LM101A), and the super-beta input (LM108). Each of Widlar's designs became a product champion in its class, undoubtedly because they all had at least one feature that was far ahead of the crowd.

13.1 The Two-Stage CMOS Op Amp

The first op-amp circuit we shall study is the two-stage CMOS topology shown in Fig. 13.1. This simple but elegant circuit has become a classic and is used in a variety of forms in the design of VLSI systems. We have already studied this circuit in Section 9.6.1 as an example of a multistage CMOS amplifier. We urge the reader to review Section 9.6.1 before proceeding further. Here, our detailed study will emphasize the performance characteristics of the circuit and the trade-offs involved in its design.

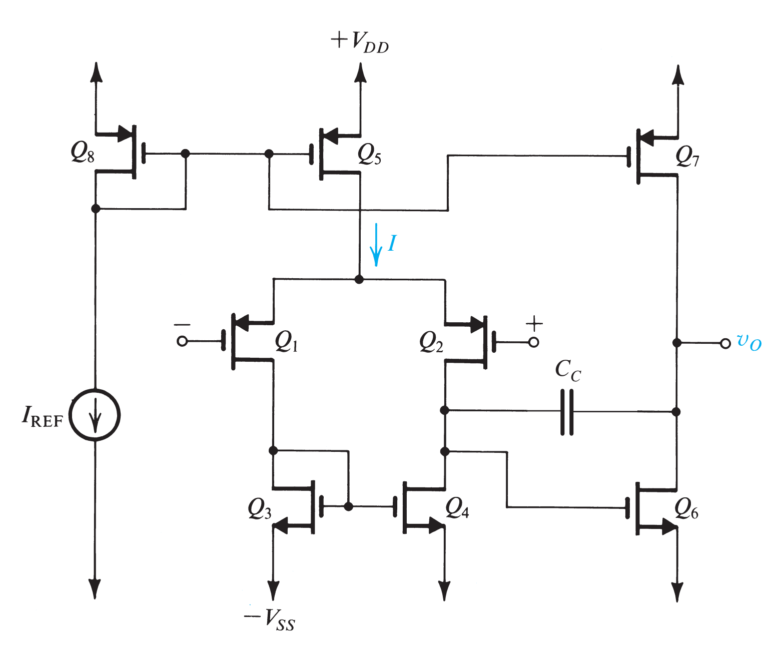


Figure 13.1 The basic two-stage CMOS op-amp configuration.

13.1.1 The Circuit

The circuit consists of two gain stages: The first stage is formed by the differential pair Q_1–Q_2 together with its current-mirror load Q_3–Q_4. This differential-amplifier circuit, studied in detail in Section 9.5, provides a voltage gain that is typically in the range of 20 V/V to 60 V/V, as well as performing conversion from differential to single-ended form while providing a reasonably high common-mode rejection ratio (CMRR).

The differential pair is biased by current source Q_5, which is one of the two output transistors of the current mirror formed by Q_8, Q_5, and Q_7. The current mirror is fed by a reference current I_{REF}, which can be generated by simply connecting a precision resistor (external to the chip) to the negative supply voltage $-V_{SS}$ or to a more precise negative voltage reference if one is available in the same integrated circuit. Alternatively, for applications with more stringent requirements, I_{REF} can be generated using a circuit such as that studied later in this section (see Fig. 13.8).

The second gain stage consists of the common-source transistor Q_6 and its current-source load Q_7. The second stage typically provides a gain of 50 V/V to 80 V/V. In addition, it takes part in the process of frequency compensating the op amp. From Section 11.10 the reader will recall that to guarantee that the op amp will operate in a stable fashion (as opposed to oscillating) when negative feedback of various amounts is applied, the open-loop gain is made to roll off with frequency at the uniform rate of –20 dB/decade. This in turn is achieved by introducing a pole at a relatively low frequency and arranging for it to dominate the frequency-response determination. In the circuit we are studying, this is implemented using a compensation capacitance C_C connected in the negative-feedback path of the second-stage amplifying transistor Q_6. As will be seen, C_C (together with the much smaller capacitance C_{gd6} across it) is Miller-multiplied by the gain of the second stage, and the resulting capacitance at the input of the second stage interacts with the total resistance there to provide the required dominant pole (more on this later).

Unless properly designed, the CMOS op-amp circuit of Fig. 13.1 can exhibit a **systematic output dc offset** voltage. This point was discussed in Section 9.6.1, where it was found that the systematic dc offset can be eliminated by sizing the transistors so as to satisfy the following constraint:

$$\frac{(W/L)_6}{(W/L)_4} = 2\frac{(W/L)_7}{(W/L)_5} \tag{13.1}$$

Finally, we observe that the CMOS op-amp circuit of Fig. 13.1 does not have an output stage. This is because it is usually required to drive only small on-chip capacitive loads.[1]

13.1.2 Input Common-Mode Range and Output Swing

Refer to Fig. 13.1 and consider the situation when the two input terminals are tied together and connected to a voltage V_{ICM}. The lowest value of V_{ICM} has to be sufficiently large to keep Q_1 and Q_2 in saturation. Thus, the lowest value of V_{ICM} should not be lower than the voltage at the drain of Q_1 ($-V_{SS} + V_{GS3} = -V_{SS} + V_{tn} + V_{OV3}$) by more than $|V_{tp}|$, thus

$$V_{ICM} \geq -V_{SS} + V_{tn} + V_{OV3} - |V_{tp}| \tag{13.2}$$

The highest value of V_{ICM} should ensure that Q_5 remains in saturation; that is, the voltage across Q_5, V_{SD5}, should not decrease below $|V_{OV5}|$. Equivalently, the voltage at the drain of Q_5 should not go higher than $V_{DD} - |V_{OV5}|$. Thus the upper limit of V_{ICM} is

$$V_{ICM} \leq V_{DD} - |V_{OV5}| - V_{SG1}$$

or equivalently

$$V_{ICM} \leq V_{DD} - |V_{OV5}| - |V_{tp}| - |V_{OV1}| \tag{13.3}$$

The expressions in Eqs. (13.2) and (13.3) can be combined to express the input common-mode range as

$$-V_{SS} + V_{OV3} + V_{tn} - |V_{tp}| \leq V_{ICM} \leq V_{DD} - |V_{tp}| - |V_{OV1}| - |V_{OV5}| \tag{13.4}$$

As expected, the overdrive voltages, which are important design parameters, subtract from the dc supply voltages, thereby reducing the input common-mode range. It follows that from a V_{ICM} range point of view, it is desirable to select the values of V_{OV} as low as possible. We observe from Eq. (13.4) that the lower limit of V_{ICM} is approximately within an overdrive voltage of $-V_{SS}$. The upper limit, however, is not as good; it is lower than V_{DD} by two overdrive voltages and a threshold voltage.

The extent of the signal swing allowed at the output of the op amp is limited at the lower end by the need to keep Q_6 saturated and at the upper end by the need to keep Q_7 saturated, thus

$$-V_{SS} + V_{OV6} \leq v_O \leq V_{DD} - |V_{OV7}| \tag{13.5}$$

Thus the output voltage can swing to within an overdrive voltage of each of the supply rails. This is a reasonably wide output swing and can be maximized by selecting values for $|V_{OV}|$ of Q_6 and Q_7 as low as possible.

[1]If the amplifier is required to drive low-resistance loads and thus a low output resistance is needed, a source follower can be connected to the output of the second stage.

An important requirement of an op-amp circuit is that it be possible for its output terminal to be connected back to its negative input terminal so that a unity-gain amplifier is obtained. For such a connection to be possible, there must be a substantial overlap between the allowable range of v_O and the allowable range of V_{ICM}. This is usually the case in the CMOS amplifier circuit under study.

EXERCISE

13.1 For a particular design of the two-stage CMOS op amp of Fig. 13.1, ±1.65-V supplies are utilized and all transistors except for Q_6 and Q_7 are operated with overdrive voltages of 0.3-V magnitude; Q_6 and Q_7 use overdrive voltages of 0.5-V magnitude. The fabrication process employed provides $V_{tn} = |V_{tp}| = 0.5$ V. Find the input common-mode range and the range allowed for v_O.
Ans. –1.35 V to 0.55 V; –1.15 V to +1.15 V

13.1.3 DC Voltage Gain

To determine the dc voltage gain and the frequency response, consider a simplified equivalent-circuit model for the small-signal operation of the CMOS amplifier (Fig. 13.2), where each of the two stages is modeled as a transconductance amplifier. As expected, the input resistance is practically infinite,

$$R_{\text{in}} = \infty$$

The first-stage transconductance G_{m1} is equal to the transconductance of each of Q_1 and Q_2 (see Section 9.5),

$$G_{m1} = g_{m1} = g_{m2} \tag{13.6}$$

Since Q_1 and Q_2 are operated at equal bias currents ($I/2$) and equal overdrive voltages, $|V_{OV1}| = |V_{OV2}|$,

$$G_{m1} = \frac{2(I/2)}{|V_{OV1}|} = \frac{I}{|V_{OV1}|} \tag{13.7}$$

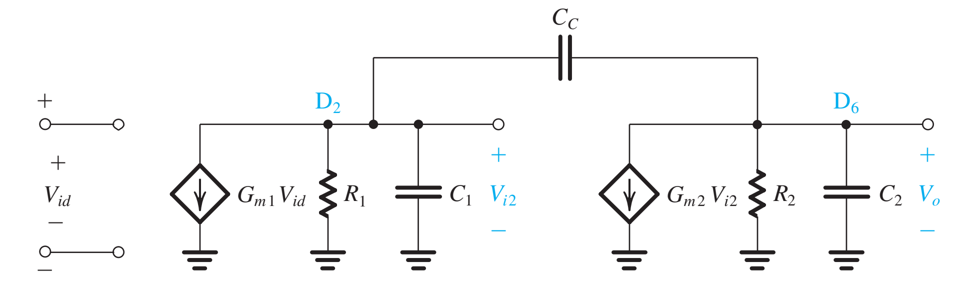


Figure 13.2 Small-signal equivalent circuit for the op amp in Fig. 13.1.

Resistance R_1 represents the output resistance of the first stage, thus

$$R_1 = r_{o2} \parallel r_{o4} \tag{13.8}$$

where

$$r_{o2} = \frac{|V_{A2}|}{I/2} \tag{13.9}$$

and

$$r_{o4} = \frac{V_{A4}}{I/2} \tag{13.10}$$

The dc gain of the first stage is thus

$$A_1 = -G_{m1}R_1 \tag{13.11}$$

$$= -g_{m1}(r_{o2} \parallel r_{o4}) \tag{13.12}$$

$$= -\frac{2}{|V_{OV1}|} \bigg/ \left[\frac{1}{|V_{A2}|} + \frac{1}{V_{A4}}\right] \tag{13.13}$$

Observe that the magnitude of A_1 is increased by operating the differential-pair transistors, Q_1 and Q_2, at a low overdrive voltage, and by choosing longer channel lengths for Q_1, Q_2, Q_3, and Q_4 so as to obtain larger Early voltages, $|V_A|$.

Returning to the equivalent circuit in Fig. 13.2 and leaving the discussion of the various model capacitances until Section 13.1.5, we note that the second-stage transconductance G_{m2} is given by

$$G_{m2} = g_{m6} = \frac{2I_{D6}}{V_{OV6}} \tag{13.14}$$

Resistance R_2 represents the output resistance of the second stage, thus

$$R_2 = r_{o6} \parallel r_{o7} \tag{13.15}$$

where

$$r_{o6} = \frac{V_{A6}}{I_{D6}} \tag{13.16}$$

and

$$r_{o7} = \frac{|V_{A7}|}{I_{D7}} = \frac{|V_{A7}|}{I_{D6}} \tag{13.17}$$

The voltage gain of the second stage can now be found as

$$A_2 = -G_{m2}R_2 \tag{13.18}$$

$$= -g_{m6}(r_{o6} \parallel r_{o7}) \tag{13.19}$$

$$= -\frac{2}{V_{OV6}} \bigg/ \left[\frac{1}{V_{A6}} + \frac{1}{|V_{A7}|}\right] \tag{13.20}$$

Here again we observe that to increase the magnitude of A_2, Q_6 has to be operated at a low overdrive voltage, and the channel lengths of Q_6 and Q_7 should be made longer.

The overall dc voltage gain can be found as the product A_1A_2,

$$A_v = A_1A_2$$

$$= G_{m1}R_1G_{m2}R_2 \quad (13.21)$$

$$= g_{m1}(r_{o2} \| r_{o4})g_{m6}(r_{o6} \| r_{o7}) \quad (13.22)$$

Note that A_v is of the order of $(g_m r_o)^2$. Thus the value of A_v will be in the range of 500 V/V to 5000 V/V.

Finally, we note that the output resistance of the op amp is equal to the output resistance of the second stage,

$$R_o = r_{o6} \| r_{o7} \quad (13.23)$$

Hence R_o can be large (i.e., in the tens-of-kilohms range). Nevertheless, as we learned from the study of negative feedback in Chapter 11, application of negative feedback that samples the op-amp output voltage results in reducing the output resistance by a factor equal to the amount of feedback $(1 + A\beta)$. Also, as mentioned before, CMOS op amps are rarely required to drive heavy resistive loads.

EXERCISES

13.2 The CMOS op amp of Fig. 13.1 is fabricated in a process for which $V'_{An} = |V'_{Ap}| = 20$ V/μm. Find A_1, A_2, and A_v if all devices are 1 μm long, $|V_{OV1}| = 0.2$ V, and $V_{OV6} = 0.5$ V. Also, find the op-amp output resistance obtained when the second stage is biased at 0.5 mA.

Ans. –100 V/V; –40 V/V; 4000 V/V; 20 kΩ

13.3 If the CMOS op amp in Fig. 13.1 is connected as a unity-gain buffer, show that the closed-loop output resistance is given by

$R_{\text{out}} \simeq 1/g_{m6}\left[g_{m1}\left(r_{o2} \| r_{o4}\right)\right]$

13.1.4 Common-Mode Rejection Ratio (CMRR)

The CMRR of the two-stage op amp of Fig. 13.1 is determined by the first stage. This was analyzed in Section 9.5.5 and the result is given in Eq. (9.158), namely,

$$\text{CMRR} = [g_{m1}(r_{o2} \| r_{o4})][2g_{m3}R_{SS}] \quad (13.24)$$

where R_{SS} is the output resistance of the bias current source $Q_5(r_{o5})$. Observe that CMRR is of the order of $(g_m r_o)^2$ and thus can be reasonably high. Also, since $g_m r_o$ is proportional to $V_A/V_{OV} = V'_A L/V_{OV}$, the CMRR is increased if long channels are used, especially for Q_5, and the transistors are operated at low overdrive voltages.

13.1.5 Frequency Response

Refer to the equivalent circuit in Fig. 13.2. Capacitance C_1 is the total capacitance between the output node of the first stage and ground, thus

$$C_1 = C_{gd2} + C_{db2} + C_{gd4} + C_{db4} + C_{gs6} \tag{13.25}$$

Capacitance C_2 represents the total capacitance between the output node of the op amp and ground and includes whatever load capacitance C_L that the amplifier is required to drive, thus

$$C_2 = C_{db6} + C_{db7} + C_{gd7} + C_L \tag{13.26}$$

Usually, C_L is larger than the transistor capacitances, with the result that C_2 becomes much larger than C_1. As mentioned before, capacitor C_C is deliberately included for the purpose of equipping the op amp with a uniform –6-dB/octave frequency response. In the following, we shall see how this is possible and how to select a value for C_C. Finally, note that in the equivalent circuit of Fig. 13.2 we should have included C_{gd6} in parallel with C_C. Usually, however, $C_C \gg C_{gd6}$, which is the reason we have neglected C_{gd6}.

To determine V_o, analysis of the circuit in Fig. 13.2 proceeds as follows. Writing a node equation at node D_2 yields

$$G_{m1}V_{id} + \frac{V_{i2}}{R_1} + sC_1V_{i2} + sC_C(V_{i2} - V_o) = 0 \tag{13.27}$$

Writing a node equation at node D_6 yields

$$G_{m2}V_{i2} + \frac{V_o}{R_2} + sC_2V_o + sC_C(V_o - V_{i2}) = 0 \tag{13.28}$$

To eliminate V_{i2} and thus determine V_o in terms of V_{id}, we use Eq. (13.28) to express V_{i2} in terms of V_o and substitute the result into Eq. (13.27). After some straightforward manipulations we obtain the amplifier transfer function

$$\frac{V_o}{V_{id}} = \frac{G_{m1}(G_{m2} - sC_C)R_1R_2}{1 + s[C_1R_1 + C_2R_2 + C_C(G_{m2}R_1R_2 + R_1 + R_2)] + s^2[C_1C_2 + C_C(C_1 + C_2)]R_1R_2} \tag{13.29}$$

First we note that for $s = 0$ (i.e., dc), Eq. (13.29) gives $V_o/V_{id} = (G_{m1}R_1)(G_{m2}R_2)$, which is what we should have expected. Second, the transfer function in Eq. (13.29) indicates that the amplifier has a transmission zero at $s = s_Z$, which is determined from

$$G_{m2} - s_ZC_C = 0$$

Thus,

$$s_Z = \frac{G_{m2}}{C_C} \tag{13.30}$$

In other words, the zero is on the positive real axis with a frequency ω_Z of

$$\omega_Z = \frac{G_{m2}}{C_C} \tag{13.31}$$

Also, the amplifier has two poles that are the roots of the denominator polynomial of Eq. (13.29). If the frequencies of the two poles are denoted ω_{P1} and ω_{P2}, then the denominator polynomial can be expressed as

$$D(s) = \left(1 + \frac{s}{\omega_{P1}}\right)\left(1 + \frac{s}{\omega_{P2}}\right) = 1 + s\left(\frac{1}{\omega_{P1}} + \frac{1}{\omega_{P2}}\right) + \frac{s^2}{\omega_{P1}\omega_{P2}}$$

Now if one of the poles is dominant, say with frequency ω_{P1}, then $\omega_{P1} \ll \omega_{P2}$, and $D(s)$ can be approximated by

$$D(s) \simeq 1 + \frac{s}{\omega_{P1}} + \frac{s^2}{\omega_{P1}\omega_{P2}} \tag{13.32}$$

The frequency of the dominant pole, ω_{P1}, can now be determined by equating the coefficients of the s terms in the denominator in Eq. (13.29) and in Eq. (13.32),

$$\begin{aligned}\omega_{P1} &= \frac{1}{C_1R_1 + C_2R_2 + C_C(G_{m2}R_2R_1 + R_1 + R_2)} \\ &= \frac{1}{R_1[C_1 + C_C(1 + G_{m2}R_2)] + R_2(C_2 + C_C)}\end{aligned} \tag{13.33}$$

We recognize the first term in the denominator as arising at the interface between the first and second stages. Here, R_1, the output resistance of the first stage, is interacting with the total capacitance at the interface. The latter is the sum of C_1 and the Miller capacitance $C_C(1 + G_{m2}R_2)$, which results from connecting C_C in the negative-feedback path of the second stage whose gain is $G_{m2}R_2$. Now, since R_1 and R_2 are usually of comparable value, we see that the first term in the denominator will be much larger than the second and we can approximate ω_{P1} as

$$\omega_{P1} \simeq \frac{1}{R_1[C_1 + C_C(1 + G_{m2}R_2)]}$$

A further approximation is possible because C_1 is usually much smaller than the Miller capacitance and $G_{m2}R_2 \gg 1$, thus

$$\omega_{P1} \simeq \frac{1}{R_1C_CG_{m2}R_2} \tag{13.34}$$

The frequency of the second, nondominant pole can be found by equating the coefficients of the s^2 terms in the denominator of Eq. (13.29) and in Eq. (13.32) and substituting for ω_{P1} from Eq. (13.34). The result is

$$\omega_{P2} = \frac{G_{m2}C_C}{C_1C_2 + C_C(C_1 + C_2)}$$

Since $C_1 \ll C_2$ and $C_1 \ll C_C$, ω_{P2} can be approximated as

$$\omega_{P2} \simeq \frac{G_{m2}}{C_2} \tag{13.35}$$

To provide the op amp with a uniform gain rolloff of –20 dB/decade down to 0 dB, the value of the compensation capacitor C_C is selected so that the resulting value of ω_{P1} (Eq. 13.34),

when multiplied by the dc gain ($G_{m1}R_1G_{m2}R_2$), results in a unity-gain frequency ω_t lower than ω_Z and ω_{P2}. Specifically

$$\omega_t = (G_{m1}R_1G_{m2}R_2)\omega_{P1} = \frac{G_{m1}}{C_C} \tag{13.36}$$

which must be lower than $\omega_Z = \dfrac{G_{m2}}{C_C}$ and $\omega_{P2} \simeq \dfrac{G_{m2}}{C_2}$. Thus, the design must satisfy the following two conditions:

$$\frac{G_{m1}}{C_C} < \frac{G_{m2}}{C_2} \tag{13.37}$$

$$G_{m1} < G_{m2} \tag{13.38}$$

EXERCISE

D13.4 Consider the frequency response of the op amp analyzed in Chapter 9 (see Example 9.6). Let $C_1 = 0.1$ pF and $C_2 = 2$ pF. Find the value of C_C that results in $f_t = 10$ MHz and verify that f_t is lower than f_Z and f_{P2}. Recall from the results of Example 9.6 that $G_{m1} = 0.3$ mA/V and $G_{m2} = 0.6$ mA/V.
Ans. $C_C = 4.8$ pF; $f_Z = 20$ MHz; $f_{P2} = 48$ MHz

Simplified Equivalent Circuit The uniform –20-dB/decade gain rolloff obtained at frequencies $f \gg f_{P1}$ but lower than f_{P2} and f_Z suggests that at these frequencies, the op amp can be represented by the simplified equivalent circuit shown in Fig. 13.3. Observe that this attractive simplification is based on the assumption that the gain of the second stage, $|A_2|$, is large, and hence a virtual ground appears at the input terminal of the second stage. The second stage then effectively acts as an integrator that is fed with the output current signal of the first stage; $G_{m1}V_{id}$. Although derived for the CMOS amplifier, this simplified equivalent circuit is general and applies to a variety of two-stage op amps, including the first two stages of the 741-type bipolar op amp studied later in this chapter.

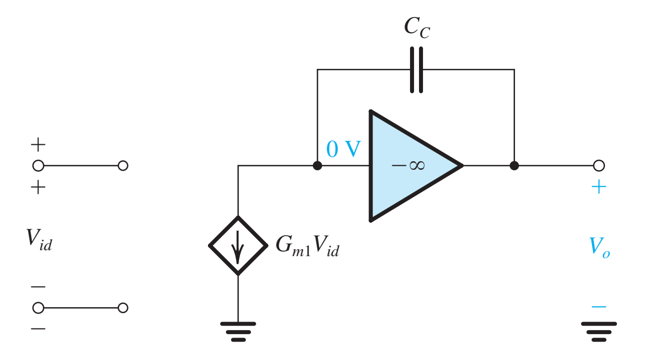


Figure 13.3 An approximate high-frequency equivalent circuit of the two-stage op amp. This circuit applies for frequencies $f \gg f_{P1}$ but lower than f_{P2} and f_Z.

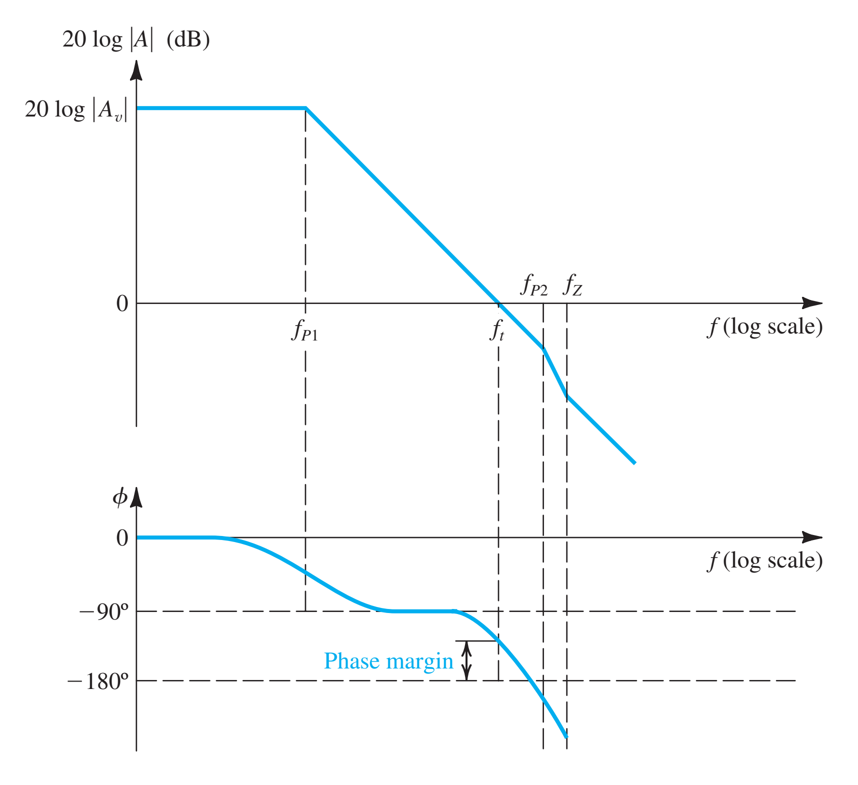


Figure 13.4 Typical frequency response of the two-stage op amp.

Phase Margin The frequency compensation scheme utilized in the two-stage CMOS amplifier is of the pole-splitting type, studied in Section 11.10.3: It provides a dominant low-frequency pole with frequency f_{P1} and shifts the second pole beyond f_t. Figure 13.4 shows a representative Bode plot for the gain magnitude and phase. Note that at the unity-gain frequency f_t, the phase lag exceeds the 90° caused by the dominant pole at f_{P1}. This so-called excess phase shift is due to the second pole,

$$\phi_{P2} = -\tan^{-1}\left(\frac{f_t}{f_{P2}}\right) \tag{13.39}$$

and the right-half-plane zero,

$$\phi_Z = -\tan^{-1}\left(\frac{f_t}{f_Z}\right) \tag{13.40}$$

Thus the phase lag at $f = f_t$ will be

$$\phi_{\text{total}} = 90° + \tan^{-1}(f_t/f_{P2}) + \tan^{-1}(f_t/f_Z) \tag{13.41}$$

and thus the phase margin will be

$$\begin{aligned}\text{Phase margin} &= 180° - \phi_{\text{total}} \\ &= 90° - \tan^{-1}(f_t/f_{P2}) - \tan^{-1}(f_t/f_Z)\end{aligned} \tag{13.42}$$

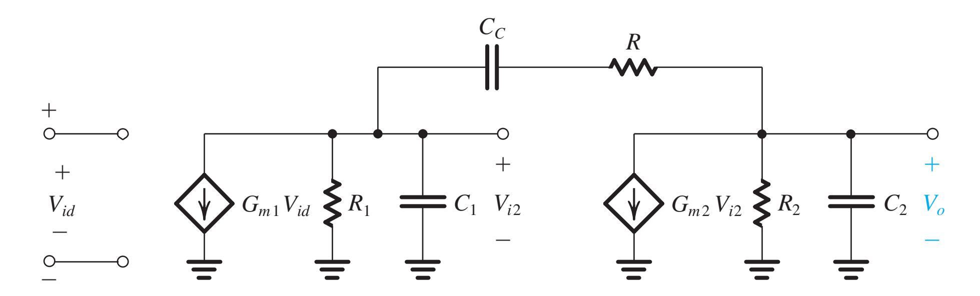


Figure 13.5 Small-signal equivalent circuit of the op amp in Fig. 13.1 with a resistance R included in series with C_C.

From our study of the stability of feedback amplifiers in Section 11.9.2, we know that the magnitude of the phase margin significantly affects the closed-loop gain.[2] Therefore, obtaining a desired minimum value of phase margin is usually a design requirement.

The problem of the additional phase lag provided by the right-half-plane zero has a rather simple and elegant solution: By including a resistance R in series with C_C, as shown in Fig. 13.5, the transmission zero can be moved to other less harmful locations. To find the new location of the transmission zero, set $V_o = 0$. Then, the current through C_C and R will be $V_{i2}/(R + 1/sC_C)$, and a node equation at the output yields

$$\frac{V_{i2}}{R + \dfrac{1}{sC_C}} = G_{m2}V_{i2} \tag{13.43}$$

Thus the zero is now at



$$s = 1\Big/C_C\left(\frac{1}{G_{m2}} - R\right) \tag{13.44}$$

We observe that by selecting $R = 1/G_{m2}$, we can place the zero at infinite frequency. An even better choice would be to select R greater than $1/G_{m2}$, thus placing the zero at a negative real-axis location where the phase it introduces becomes a phase lead and thus *adds* to the phase margin.

EXERCISE

13.5 A particular implementation of the CMOS amplifier of Figs. 13.1 and 13.2 provides $G_{m1} = 1$ mA/V, $G_{m2} = 2$ mA/V, $r_{o2} = r_{o4} = 100$ kΩ, $r_{o6} = r_{o7} = 40$ kΩ, and $C_2 = 1$ pF.
(a) Find the value of C_C that results in $f_t = 100$ MHz. What is the 3-dB frequency of the open-loop gain?
(b) Find the value of the resistance R that when placed in series with C_C causes the transmission zero to be located at infinite frequency.
(c) Find the frequency of the second pole and hence find the excess phase lag at $f = f_t$, introduced by the second pole, and the resulting phase margin assuming that the situation in (b) pertains.

Ans. 1.6 pF; 50 kHz; 500 Ω; 318 MHz; 17.4°; 72.6°

[2]The magnitude of the phase margin also affects the step response of the closed-loop amplifier.

13.1.6 Slew Rate

The slew-rate limitation of op amps is discussed in Chapter 2. Here, we shall illustrate the origin of the slewing phenomenon in the context of the two-stage CMOS amplifier under study.

Consider the unity-gain follower of Fig. 13.6 with a step of, say, 1 V applied at the input. Because of the amplifier dynamics, its output will not change in zero time. Thus, immediately after the input is applied, the entire value of the step will appear as a differential signal between the two input terminals. In all likelihood, such a large signal will exceed the voltage required to turn off one side of the input differential pair ($\sqrt{2}V_{OV1}$: see earlier illustration in Chapter 9, Fig. 9.6) and switch the entire bias current I to the other side. Reference to Fig. 13.1 shows that for our example, Q_2 will turn off, and Q_1 will conduct the entire current I. Thus Q_4 will sink a current I that will be pulled from C_C, as shown in Fig. 13.7. Here, as we did in Fig. 13.3, we are modeling the second stage as an ideal integrator. We see that the output voltage will be a ramp with a slope of I/C_C:

$$v_O(t) = \frac{I}{C_C}t$$

Thus the slew rate, SR, is given by

$$SR = \frac{I}{C_C} \tag{13.45}$$

It should be pointed out, however, that this is a rather simplified model of the slewing process.

Relationship Between SR and f_t A simple relationship exists between the unity-gain bandwidth f_t and the slew rate SR. This relationship can be found by combining Eqs. (13.36) and (13.45) and noting that $G_{m1} = g_{m1} = I/V_{OV1}$, to obtain

$$SR = 2\pi f_t V_{OV1} \tag{13.46}$$

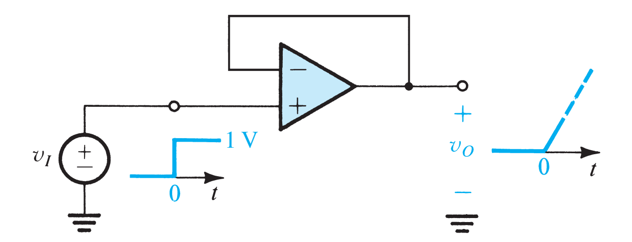


Figure 13.6 A unity-gain follower with a large step input. Since the output voltage cannot change immediately, a large differential voltage appears between the op-amp input terminals.

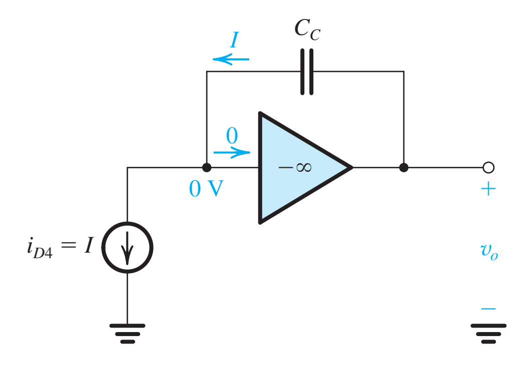


Figure 13.7 Model of the two-stage CMOS op-amp of Fig. 13.1 when a large differential voltage is applied.

or equivalently,



$$SR = V_{OV1}\omega_t \tag{13.47}$$

Thus, for a given ω_t, the slew rate is determined by the overdrive voltage at which the first-stage transistors are operated. A higher slew rate is obtained by operating Q_1 and Q_2 at a larger V_{OV}. Now, for a given bias current I, a larger V_{OV} is obtained if Q_1 and Q_2 are p-channel devices. This is an important reason for using p-channel rather than n-channel devices in the first stage of the CMOS op amp. Another reason is that it allows the second stage to employ an n-channel device. Now, since n-channel devices have greater transconductances than corresponding p-channel devices, G_{m2} will be high, resulting in a higher second-pole frequency and a correspondingly higher ω_t. However, the price paid for these improvements is a lower G_{m1} and hence a lower dc gain.

EXERCISE

13.6 Find SR for the CMOS op amp of Fig. 13.1 for the case $f_t = 100$ MHz and $V_{OV1} = 0.2$ V. If $C_C = 1.6$ pF, what must the bias current I be?
Ans. 126 V/μs; 200 μA

13.1.7 Power-Supply Rejection Ratio (PSRR)

CMOS op amps are usually utilized in what are known as **mixed-signal circuits**: IC chips that combine analog and digital circuits. In such circuits, the switching activity in the digital portion usually results in increased ripple on the power supplies. A portion of the supply ripple can make its way to the op-amp output and thus corrupt the output signal. The traditional approach for reducing supply ripple by connecting large capacitances between the supply rails and ground is not viable in IC design, as such capacitances would consume most of the chip area. Instead, the analog IC designer has to pay attention to another op-amp specification that so far we have ignored, namely, the power-supply rejection ratio (PSRR).

The PSRR is defined as the ratio of the amplifier differential gain to the gain experienced by a change in the power-supply voltage (v_{dd} and v_{ss}). For circuits utilizing two power supplies, we define

$$\text{PSRR}^+ \equiv \frac{A_d}{A^+} \tag{13.48}$$

and

$$\text{PSRR}^- = \frac{A_d}{A^-} \tag{13.49}$$

where

$$A^+ \equiv \frac{v_o}{v_{dd}} \tag{13.50}$$

$$A^- = \frac{v_o}{v_{ss}} \tag{13.51}$$

Obviously, to minimize the effect of the power-supply ripple, we require the op amp to have a large PSRR.

A detailed analysis of the PSRR of the two-stage CMOS op amp is beyond the scope of this book (see Gray et al., 2009). Nevertheless, we make the following brief remarks. It can be shown that the circuit is remarkably insensitive to variations in V_{DD}, and thus PSRR$^+$ is very high. This is not the case, however, for the negative-supply ripple v_{ss}, which is coupled to the output primarily through the second-stage transistors Q_6 and Q_7. In particular, the portion of v_{ss} that appears at the op-amp output is determined by the voltage divider formed by the output resistances of Q_6 and Q_7,

$$v_o = v_{ss}\frac{r_{o7}}{r_{o6}+r_{o7}} \quad (13.52)$$

Thus,

$$A^- \equiv \frac{v_o}{v_{ss}} = \frac{r_{o7}}{r_{o6}+r_{o7}} \quad (13.53)$$

Now utilizing A_d from Eq. (13.22) gives

$$\text{PSRR}^- \equiv \frac{A_d}{A^-} = g_{m1}(r_{o2} \parallel r_{o4})g_{m6}r_{o6} \quad (13.54)$$

Thus, PSRR$^-$ is of the form $(g_m r_o)^2$ and therefore is maximized by selecting long channels L (to increase $|V_A|$), and operating at low $|V_{OV}|$.

13.1.8 Design Trade-Offs

The performance parameters of the two-stage CMOS amplifier are primarily determined by two design parameters:

1. The length L used for the channel of each MOSFET.
2. The overdrive voltage $|V_{OV}|$ at which each transistor is operated.

Throughout this section, we have found that a larger L and correspondingly larger $|V_A|$ increases the amplifier gain, CMRR and PSRR. We also found that operating at a lower $|V_{OV}|$ increases these three parameters as well as increasing the input common-mode range and the allowable range of output swing. Also, although we have not analyzed the offset voltage of the op amp here, we know from our study of the subject in Section 9.4.1 that a number of the components of the input offset voltage that arises from random device mismatches are proportional to $|V_{OV}|$ at which the MOSFETs of the input differential pair are operated. Thus the offset is minimized by operating at a lower $|V_{OV}|$.

There is, however, an important MOSFET performance parameter that requires the selection of a larger $|V_{OV}|$, namely, **the transition frequency** f_T, which determines the high-frequency performance of the MOSFET (see Section 10.2.1),

$$f_T = \frac{g_m}{2\pi\left(C_{gs}+C_{gd}\right)} \quad (13.55)$$

For an n-channel MOSFET, it can be shown that[3]

$$f_T \simeq \frac{1.5\mu_n V_{OV}}{2\pi L^2} \tag{13.56}$$

A similar relationship applies for the PMOS transistor, with μ_p and $|V_{OV}|$ replacing μ_n and V_{OV}, respectively. Thus to increase f_T and improve the high-frequency response of the op amp, we need to use a larger overdrive value and, not surprisingly, shorter channels. A larger $|V_{OV}|$ also results in a higher op-amp slew rate SR (Eq. 13.46). Finally, note that the selection of a larger $|V_{OV}|$ results, for the same bias current, and thus the same power dissipation, in a smaller W/L, which combined with a short L leads to smaller devices and hence lower values of MOSFET capacitances and higher frequencies of operation.

In conclusion, the selection of $|V_{OV}|$ presents the designer with a trade-off between improving the low-frequency performance parameters on the one hand and the high-frequency performance on the other. For modern submicron technologies, which require operation from power supplies of 1 V to 1.5 V, overdrive voltages between 0.1 V and 0.3 V are typically utilized. For these process technologies, analog designers typically use channel lengths that are at least 1.5 to 2 times the specified value of $L_{\min}$, and even longer channels are used for current-source bias transistors.

13.1.9 A Bias Circuit for the Two-Stage CMOS Op Amp

We now present a circuit for generating the bias current I_{REF} of the two-stage CMOS op amp of Fig. 13.1. As will be seen, the value of the bias current generated is independent of both the supply voltage and the threshold voltage of the MOSFETs. As well, the transconductance of each MOSFET biased by this circuit (i.e., by a multiple of I_{REF}) has a value determined by only a single resistor and the device dimensions.

The bias circuit is shown in Fig. 13.8. It consists of two deliberately mismatched transistors, Q_{12} and Q_{13}, with Q_{12} usually about four times wider than Q_{13}. A resistor R_B is connected in series with the source of Q_{12}. Since, as will be shown, R_B determines both

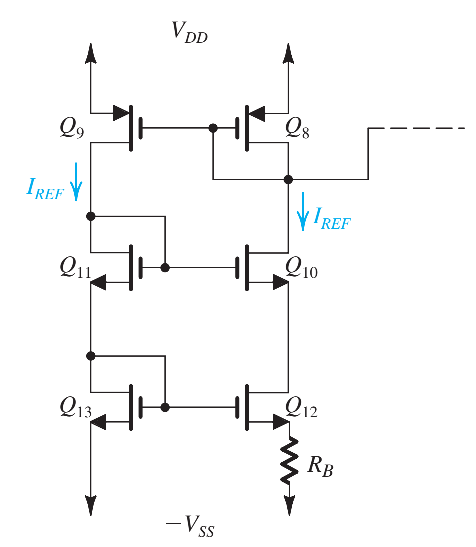


Figure 13.8 Bias circuit for the CMOS op amp. Note that Q_8 is the same Q_8 in the circuit of Fig. 13.1.

[3] See Appendix G on the companion website.

the bias current I_{REF} and the transconductance g_{m12}, its value should be accurate and stable; in most applications, R_B would be an off-chip resistor. In order to minimize the channel-length modulation effect on Q_{12}, we include a cascode transistor Q_{10} and a matched diode-connected transistor Q_{11} to provide a bias voltage for Q_{10}. Finally, a *p*-channel current mirror formed by a pair of matched devices, Q_8 and Q_9, both replicates the current I_B back to Q_{11} and Q_{13} and provides a bias line for Q_5 and Q_7 of the CMOS op-amp circuit of Fig. 13.11.

The circuit operates as follows: The current mirror (Q_8, Q_9) causes Q_{13} to conduct a current equal to that in Q_{12}, that is, I_{REF}. Thus,

$$I_{REF} = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_{12}(V_{GS12} - V_t)^2 \tag{13.57}$$

and,

$$I_{REF} = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_{13}(V_{GS13} - V_t)^2 \tag{13.58}$$

From the circuit, we see that the gate-source voltages of Q_{12} and Q_{13} are related by

$$V_{GS13} = V_{GS12} + I_{REF}R_B$$

Subtracting V_t from both sides of this equation and using Eqs. (13.57) and (13.58) to replace $(V_{GS12} - V_t)$ and $(V_{GS13} - V_t)$ results in

$$\sqrt{\frac{2I_{REF}}{\mu_n C_{ox}(W/L)_{13}}} = \sqrt{\frac{2I_{REF}}{\mu_n C_{ox}(W/L)_{12}}} + I_{REF}R_B \tag{13.59}$$

This equation can be rearranged to yield

$$I_{REF} = \frac{2}{\mu_n C_{ox}(W/L)_{12}R_B^2}\left(\sqrt{\frac{(W/L)_{12}}{(W/L)_{13}}} - 1\right)^2 \tag{13.60}$$

from which we observe that I_{REF} is determined by the dimensions of Q_{12} and the value of R_B and by the ratio of the dimensions of Q_{12} and Q_{13}. Furthermore, Eq. (13.60) can be rearranged to the form

$$R_B = \frac{2}{\sqrt{2\mu_n C_{ox}(W/L)_{12}I_{REF}}}\left(\sqrt{\frac{(W/L)_{12}}{(W/L)_{13}}} - 1\right)$$

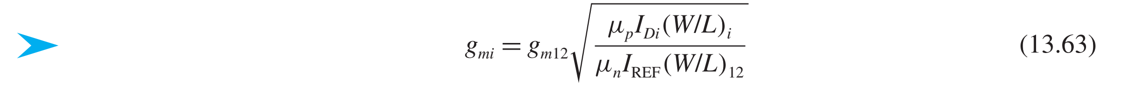
in which we recognize the factor $\sqrt{\mu_n C_{ox}(W/L)_{12}I_{REF}}$ as g_{m12}; thus,

$$g_{m12} = \frac{2}{R_B}\left(\sqrt{\frac{(W/L)_{12}}{(W/L)_{13}}} - 1\right) \tag{13.61}$$

This is a very interesting result: g_{m12} is determined solely by the value of R_B and the ratio of the dimensions of Q_{12} and Q_{13}. Furthermore, since g_m of a MOSFET is proportional to $\sqrt{I_D(W/L)}$, each transistor biased by the circuit of Fig. 13.8; that is, each transistor whose bias current is derived from I_{REF} will have a g_m value that is a multiple of g_{m12}. Specifically, the *i*th *n*-channel MOSFET will have

$$g_{mi} = g_{m12}\sqrt{\frac{I_{Di}(W/L)_i}{I_{REF}(W/L)_{12}}} \tag{13.62}$$

and the ith p-channel device will have



$$g_{mi} = g_{m12}\sqrt{\frac{\mu_p I_{Di}(W/L)_i}{\mu_n I_{\text{REF}}(W/L)_{12}}} \tag{13.63}$$

Finally, it should be noted that the bias circuit of Fig. 13.8 employs **positive feedback**, and thus care should be exercised in its design to avoid unstable performance. Instability is avoided by making Q_{12} wider than Q_{13}, as has already been pointed out. Nevertheless, some form of instability may still occur; in fact, the circuit can operate in a stable state in which all currents are zero. To get it out of this state, current needs to be injected into one of its nodes, to "kick start" its operation. Feedback and stability are studied in Chapter 11.

EXERCISES

13.7 Consider the bias circuit of Fig. 13.8 for the case of $(W/L)_8 = (W/L)_9 = (W/L)_{10} = (W/L)_{11} = (W/L)_{13} = 20$ and $(W/L)_{12} = 80$. The circuit is fabricated in a process technology for which $\mu_n C_{ox} = 90\ \mu\text{A/V}^2$. Find the value of R_B that results in a bias current $I_{\text{REF}} = 10\ \mu\text{A}$. Also, find the transconductance g_{m12}.
Ans. 5.27 kΩ; 0.379 mA/V

D13.8 Design the bias circuit of Fig. 13.8 to operate with the CMOS op amp of Example 9.6. Use Q_8 and Q_9 as identical devices with Q_8 having the dimensions given in Example 9.6. Transistors Q_{10}, Q_{11}, and Q_{13} are to be identical, with the same g_m as Q_8 and Q_9. Transistor Q_{12} is to be four times as wide as Q_{13}. Find the required value of R_B. What is the voltage drop across R_B? Also give the values of the dc voltages at the gates of Q_{12}, Q_{10}, and Q_8.
Ans. 1.67 kΩ; 150 mV; −1.5 V; −0.5 V; +1.4 V

Example 13.1

We conclude our study of the two-stage CMOS op amp with a design example. Let it be required to design the circuit to obtain a dc gain of 4000 V/V. Assume that the available fabrication technology is of the 0.5-μm type for which $V_{tn} = |V_{tp}| = 0.5$ V, $k_n' = 200\ \mu\text{A/V}^2$, $k_p' = 80\ \mu\text{A/V}^2$, $V_{An}' = |V_{Ap}'| = 20$ V/μm, and $V_{DD} = V_{SS} = 1.65$ V. To achieve a reasonable dc gain per stage, use $L = 1$ μm for all devices. Also, for simplicity, operate all devices at the same $|V_{OV}|$, in the range of 0.2 V to 0.4 V. Use $I = 200\ \mu$A, and to obtain a higher G_{m2}, and hence a higher f_{P2}, use $I_{D6} = 0.5$ mA. Specify the W/L ratios for all transistors. Also give the values realized for the input common-mode range, the maximum possible output swing, R_{in} and R_o. Also determine the CMRR and PSRR realized. If $C_1 = 0.2$ pF and $C_2 = 0.8$ pF, find the required values of C_C and the series resistance R to place the transmission zero at $s = \infty$ and to obtain the highest possible f_t consistent with a phase margin of 85°. Evaluate the values obtained for f_t and SR.

Solution

Using the voltage-gain expression in Eq. (13.22),

$$A_v = g_{m1}(r_{o2} \| r_{o4}) g_{m6}(r_{o6} \| r_{o7})$$

$$= \frac{2(I/2)}{V_{OV}} \times \frac{1}{2} \times \frac{V_A}{(I/2)} \times \frac{2I_{D6}}{V_{OV}} \times \frac{1}{2} \times \frac{V_A}{I_{D6}}$$

$$= \left(\frac{V_A}{V_{OV}}\right)^2$$

To obtain $A_v = 4000$, given $V_A = 20$ V,

$$4000 = \frac{400}{V_{OV}^2}$$

$$V_{OV} = 0.316 \text{ V}$$

To obtain the required (W/L) ratios of Q_1 and Q_2,

$$I_{D1} = \frac{1}{2} k_p' \left(\frac{W}{L}\right)_1 V_{OV}^2$$

$$100 = \frac{1}{2} \times 80 \left(\frac{W}{L}\right)_1 \times 0.316^2$$

Thus,

$$\left(\frac{W}{L}\right)_1 = \frac{25 \text{ μm}}{1 \text{ μm}}$$

and

$$\left(\frac{W}{L}\right)_2 = \frac{25 \text{ μm}}{1 \text{ μm}}$$

For Q_3 and Q_4 we write

$$100 = \frac{1}{2} \times 200 \left(\frac{W}{L}\right)_3 \times 0.316^2$$

to obtain

$$\left(\frac{W}{L}\right)_3 = \left(\frac{W}{L}\right)_4 = \frac{10 \text{ μm}}{1 \text{ μm}}$$

For Q_5,

$$200 = \frac{1}{2} \times 80 \left(\frac{W}{L}\right)_5 \times 0.316^2$$

Example 13.1 *continued*

Thus,

$$\left(\frac{W}{L}\right)_5 = \frac{50\ \mu\text{m}}{1\ \mu\text{m}}$$

Since Q_7 is required to conduct 500 μA, its (W/L) ratio should be 2.5 times that of Q_5,

$$\left(\frac{W}{L}\right)_7 = 2.5\left(\frac{W}{L}\right)_5 = \frac{125\ \mu\text{m}}{1\ \mu\text{m}}$$

For Q_6 we write

$$500 = \frac{1}{2} \times 200 \times \left(\frac{W}{L}\right)_6 \times 0.316^2$$

Thus,

$$\left(\frac{W}{L}\right)_6 = \frac{50\ \mu\text{m}}{1\ \mu\text{m}}$$

At this point we should check that condition (13.1) is satisfied, which is indeed the case, ensuring that there will be no systematic output offset voltage.

Finally, let's select $I_{\text{REF}} = 20\ \mu\text{A}$, thus

$$\left(\frac{W}{L}\right)_8 = 0.1\left(\frac{W}{L}\right)_5 = \frac{5\ \mu\text{m}}{1\ \mu\text{m}}$$

The input common-mode range can be found using the expression in Eq. (13.4) as

$$-1.33\ \text{V} \le V_{ICM} \le 0.52\ \text{V}$$

The maximum signal swing allowable at the output is found using the expression in Eq. (13.5) as

$$-1.33\ \text{V} \le v_O \le 1.33\ \text{V}$$

The input resistance is practically infinite, and the output resistance is

$$R_o = r_{o6} \,\|\, r_{o7} = \frac{1}{2} \times \frac{20}{0.5} = 20\ \text{k}\Omega$$

The CMRR is determined using Eq. (13.24),

$$\text{CMRR} = g_{m1}(r_{o2} \,\|\, r_{o4})(2g_{m3}R_{SS})$$

where $R_{SS} = r_{o5} = V_A/I$. Thus,

$$\text{CMRR} = \frac{2(I/2)}{V_{OV}} \times \frac{1}{2} \times \frac{V_A}{(I/2)} \times 2 \times \frac{2(I/2)}{V_{OV}} \times \frac{V_A}{I}$$

$$= 2\left(\frac{V_A}{V_{OV}}\right)^2 = 2\left(\frac{20}{0.316}\right)^2 = 8000$$

Expressed in decibels, we have

$$\text{CMRR} = 20 \log 8000 = 78 \text{ dB}$$

The PSRR is determined using Eq. (13.53):

$$\begin{aligned}\text{PSRR} &= g_{m1}(r_{o2} \parallel r_{o4}) g_{m6} r_{o6} \\ &= \frac{2(I/2)}{V_{OV}} \times \frac{1}{2} \times \frac{V_A}{(I/2)} \times \frac{2I_{D6}}{V_{OV}} \times \frac{V_A}{I_{D6}} \\ &= 2\left(\frac{V_A}{V_{OV}}\right)^2 = 2\left(\frac{20}{0.316}\right)^2 = 8000\end{aligned}$$

or, expressed in decibels,

$$\text{PSRR} = 20 \log 8000 = 78 \text{ dB}$$

To determine f_{P2} we use Eq. (13.35) and substitute for G_{m2},

$$G_{m2} = g_{m6} = \frac{2I_{D6}}{V_{OV}} = \frac{2 \times 0.5}{0.316} = 3.2 \text{ mA/V}$$

Thus,

$$f_{P2} = \frac{3.2 \times 10^{-3}}{2\pi \times 0.8 \times 10^{-12}} = 637 \text{ MHz}$$

To move the transmission zero to $s = \infty$, we select the value of R as

$$R = \frac{1}{G_{m2}} = \frac{1}{3.2 \times 10^{-3}} = 316 \ \Omega$$

For a phase margin of 85°, the phase shift due to the second pole at $f = f_t$ must be 5°, that is,

$$\tan^{-1} \frac{f_t}{f_{P2}} = 5°$$

Thus,

$$f_t = 637 \times \tan 5° = 55.7 \text{ MHz}$$

The value of C_C can be found using Eq. (13.36),

$$C_C = \frac{G_{m1}}{2\pi f_t}$$

Example 13.1 *continued*

where

$$G_{m1} = g_{m1} = \frac{2 \times 100\ \mu\text{A}}{0.316\ \text{V}} = 0.63\ \text{mA/V}$$

Thus,

$$C_{C1} = \frac{0.63 \times 10^{-3}}{2\pi \times 55.7 \times 10^{6}} = 1.8\ \text{pF}$$

The value of *SR* can now be found using Eq. (13.46) as

$$\begin{aligned} SR &= 2\pi \times 55.7 \times 10^{6} \times 0.316 \\ &= 111\ \text{V}/\mu\text{s} \end{aligned}$$

13.2 The Folded-Cascode CMOS Op Amp

In this section we study another type of CMOS op-amp circuit: the folded cascode. The circuit is based on the folded-cascode amplifier studied in Section 8.5.5. There, it was mentioned that although composed of a CS transistor and a CG transistor of opposite polarity, the folded-cascode configuration is generally considered to be a single-stage amplifier. Similarly, the op-amp circuit that is based on the cascode configuration is considered to be a single-stage op amp. Nevertheless, it can be designed to provide performance parameters that equal and in some respects exceed those of the two-stage topology studied in the preceding section. Indeed, the folded-cascode op-amp topology is currently as popular as the two-stage structure.

13.2.1 The Circuit

Figure 13.9 shows the structure of the CMOS folded-cascode op amp. Here, Q_1 and Q_2 form the input differential pair, and Q_3 and Q_4 are the cascode transistors. Recall that for differential input signals, each of Q_1 and Q_2 acts as a common-source amplifier. Also note that the gate terminals of Q_3 and Q_4 are connected to a constant dc voltage (V_{BIAS1}) and hence are at signal ground. Thus, for differential input signals, each of the transistor pairs Q_1–Q_3 and Q_2–Q_4 acts as a folded-cascode amplifier, such as the one in Fig. 8.36. Note that the input differential pair is biased by a constant-current source I. Thus each of Q_1 and Q_2 is operating at a bias current $I/2$. A node equation at each of their drains shows that the bias current of each of Q_3 and Q_4 is $(I_B - I/2)$. As will be seen shortly, both the dc gain and the unity-gain frequency are proportional to g_m of each of Q_1 and Q_2. Thus, the bias current I is usually made large to obtain a high value for $g_{m1,2}$. For a given power dissipation and thus a given total current $2I_B$, the current that biases each of Q_3 and Q_4 $(I_B - \frac{I}{2})$ will of necessity be small. It turns out, however, that this is advantageous, as it results in a large r_o for Q_4 and thus a large output resistance and a correspondingly large dc gain for the op amp. As a rule of thumb, the ratio of $I_{D1,2}$ to $I_{D3,4}$ can be selected as large as 4.

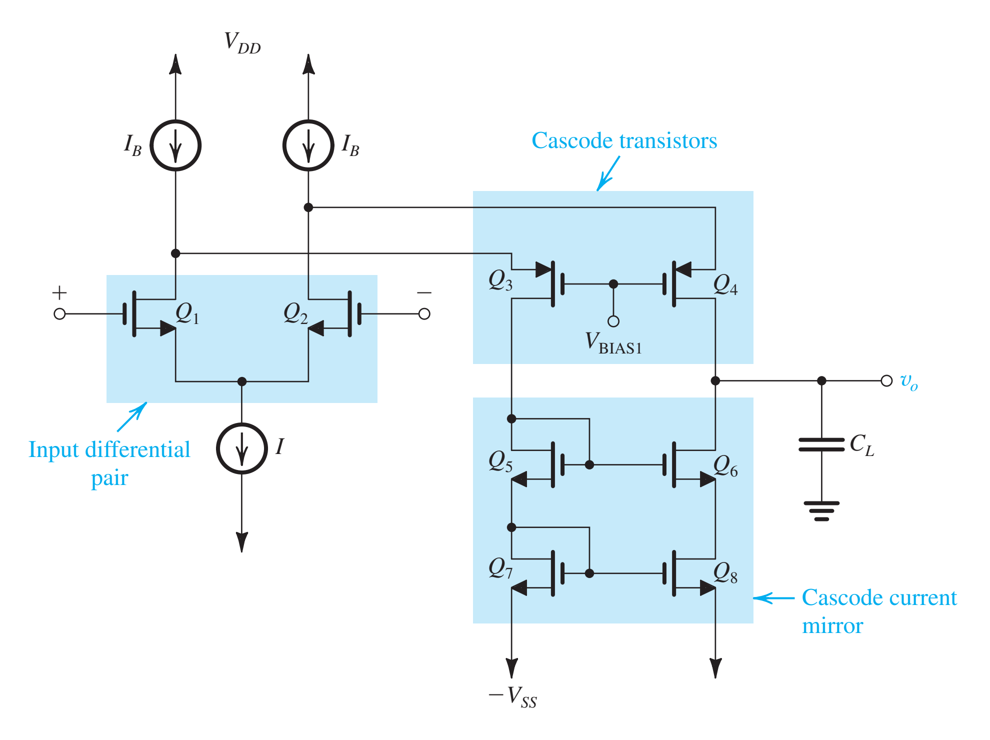


Figure 13.9 Structure of the folded-cascode CMOS op amp.

EXERCISE

13.9 To limit the power dissipation in the op-amp circuit of Fig. 13.9 to an acceptable level, the total dc current is limited to 0.3 mA. If it is desired to bias each of Q_1 and Q_2 at a dc current four times the bias current of each of Q_3 and Q_4, find the values of I_B, I, $I_{D1,2}$, and $I_{D3,4}$.
Ans. 150 μA; 240 μA; 120 μA; 30 μA

As we learned in Chapter 8, if the full advantage of the high output resistance achieved through cascoding is to be realized, the output resistance of the current-source load must be equally high. This is the reason for using the cascode current mirror Q_5 to Q_8 in the circuit of Fig. 13.9. (This current-mirror circuit was studied in Section 8.6.1.) Finally, note that capacitance C_L denotes the total capacitance at the output node. It includes the internal transistor capacitances, an actual load capacitance (if any), and possibly an additional capacitance deliberately introduced for the purpose of frequency compensation. In many cases, however, the load capacitance will be sufficiently large, obviating the need to provide additional capacitance to achieve the desired frequency compensation. This topic will be discussed shortly. For the time being, we note that unlike the two-stage circuit, which requires

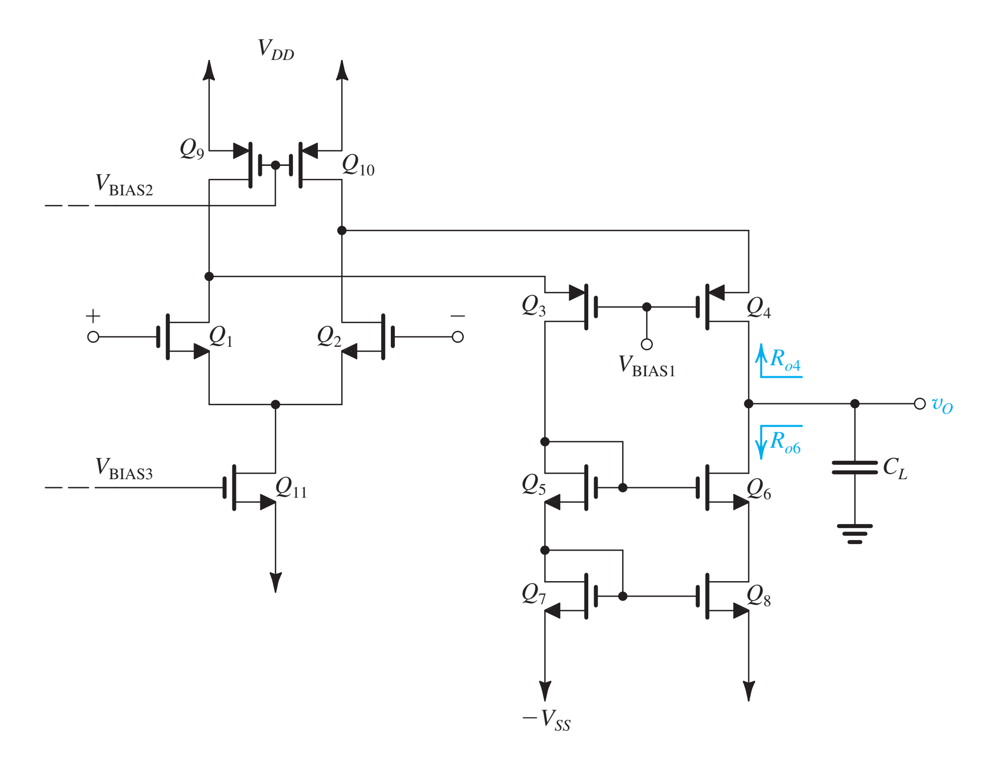


Figure 13.10 A more complete circuit for the folded-cascode CMOS amplifier of Fig. 13.9.

the introduction of a separate compensation capacitor C_C, here the load capacitance contributes to frequency compensation.

A more complete circuit for the CMOS folded-cascode op amp is shown in Fig. 13.10. Here we show the two transistors Q_9 and Q_{10}, which provide the constant bias currents I_B, and transistor Q_{11}, which provides the constant current I utilized for biasing the differential pair. Observe that the details for generating the bias voltages V_{BIAS1}, V_{BIAS2}, and V_{BIAS3} are not shown. Nevertheless, we are interested in how the values of these voltages are to be selected. Toward that end, we evaluate the input common-mode range and the allowable output swing.

13.2.2 Input Common-Mode Range and Output Swing

To find the input common-mode range, let the two input terminals be tied together and connected to a voltage V_{ICM}. The maximum value of V_{ICM} is limited by the requirement that Q_1 and Q_2 operate in saturation at all times. Thus $V_{ICM\text{max}}$ should be at most V_{tn} volts above the voltage at the drains of Q_1 and Q_2. The latter voltage is determined by V_{BIAS1} and must allow for a voltage drop across Q_9 and Q_{10} at least equal to their overdrive voltage, $|V_{OV9}| = |V_{OV10}|$. Assuming that Q_9 and Q_{10} are indeed operated at the edge of saturation, $V_{ICM\text{max}}$ will be

$$V_{ICM\text{max}} = V_{DD} - |V_{OV9}| + V_{tn} \tag{13.64}$$

which can be larger than V_{DD}, a significant improvement over the case of the two-stage circuit. The value of V_{BIAS2} should be selected to yield the required value of I_B while operating Q_9 and Q_{10} at a small value of $|V_{OV}|$ (e.g., 0.2 V or so). The minimum value of V_{ICM} is limited by the need to keep Q_{11} operating in saturation at all times, which is assured by keeping the voltage

across it no smaller than V_{OV11} at all times. Thus

$$V_{ICM\min} = -V_{SS} + V_{OV11} + V_{OV1} + V_{tn} \tag{13.65}$$

The presence of the threshold voltage V_{tn} in this expression indicates that $V_{ICM\min}$ is not sufficiently low. Later in this section we shall describe an ingenious technique for solving this problem. For the time being, note that the value of V_{BIAS3} should be selected to provide the required value of I while operating Q_{11} at a low overdrive voltage. Combining Eqs. (13.64) and (13.65) provides

$$-V_{SS} + V_{OV11} + V_{OV1} + V_{tn} \leq V_{ICM} \leq V_{DD} - |V_{OV9}| + V_{tn} \tag{13.66}$$

The upper end of the allowable range of v_O is determined by the need to maintain Q_{10} and Q_4 in saturation. Note that Q_{10} will operate in saturation as long as an overdrive voltage, $|V_{OV10}|$, appears across it. It follows that to maximize the allowable positive swing of v_O (and also $V_{ICM\max}$), we should select the value of V_{BIAS1} so that Q_{10} operates at the edge of saturation, that is,

$$V_{\text{BIAS1}} = V_{DD} - |V_{OV10}| - V_{SG4} \tag{13.67}$$

The upper limit of v_O will then be

$$v_{O\max} = V_{DD} - |V_{OV10}| - |V_{OV4}| \tag{13.68}$$

which is two overdrive voltages below V_{DD}. The situation is not as good, however, at the other end: Since the voltage at the gate of Q_6 is $-V_{SS} + V_{GS7} + V_{GS5}$ or equivalently $-V_{SS} + V_{OV7} + V_{OV5} + 2V_{tn}$, the lowest possible v_O is obtained when Q_6 reaches the edge of saturation, namely, when v_O decreases below the voltage at the gate of Q_6 by V_{tn}, that is,

$$v_{O\min} = -V_{SS} + V_{OV7} + V_{OV5} + V_{tn} \tag{13.69}$$

Note that this value is two overdrive voltages *plus* a threshold voltage above $-V_{SS}$. This is a drawback of utilizing the cascode mirror. The problem can be alleviated by using a modified mirror circuit, as we shall shortly see.

EXERCISE

13.10 For a particular design of the folded-cascode op amp of Fig. 13.10, ±1.65-V supplies are utilized and all transistors are operated at overdrive voltages of 0.3-V magnitude. The fabrication process employed provides $V_{tn} = |V_{tp}| = 0.5$ V. Find the input common-mode range and the range allowed for v_O.

Ans. –0.55 V to +1.85 V; –0.55 V to +1.05 V

13.2.3 Voltage Gain

The folded-cascode op amp is simply a transconductance amplifier with an infinite input resistance, a transconductance G_m, and an output resistance R_o. In addition, G_m is equal to g_m of each of the two transistors of the differential pair,

$$G_m = g_{m1} = g_{m2} \tag{13.70}$$

Thus,

$$G_m = \frac{2(I/2)}{V_{OV1}} = \frac{I}{V_{OV1}} \tag{13.71}$$

The output resistance R_o is the parallel equivalent of the output resistance of the cascode amplifier and the output resistance of the cascode mirror, thus

$$R_o = R_{o4} \parallel R_{o6} \tag{13.72}$$

Reference to Fig. 13.10 shows that the resistance R_{o4} is the output resistance of the CG transistor Q_4. The latter has a resistance $(r_{o2} \parallel r_{o10})$ in its source lead, thus

$$R_{o4} \simeq (g_{m4}r_{o4})(r_{o2} \parallel r_{o10}) \tag{13.73}$$

The resistance R_{o6} is the output resistance of the cascode mirror and is thus given by Eq. (8.94), thus

$$R_{o6} \simeq g_{m6}r_{o6}r_{o8} \tag{13.74}$$

Combining Eqs. (13.72) to (13.74) gives

$$R_o = [g_{m4}r_{o4}(r_{o2} \parallel r_{o10})] \parallel (g_{m6}r_{o6}r_{o8}) \tag{13.75}$$

The dc open-loop gain can now be found using G_m and R_o, as

$$A_v = G_m R_o \tag{13.76}$$

Thus,

$$A_v = g_{m1}\{[g_{m4}r_{o4}(r_{o2} \parallel r_{o10})] \parallel (g_{m6}r_{o6}r_{o8})\} \tag{13.77}$$

Figure 13.11 shows the equivalent-circuit model including the load capacitance C_L, which we shall take into account shortly.

Because the folded-cascode op amp is a transconductance amplifier, it has been given the name **operational transconductance amplifier** (**OTA**). Its very high output resistance, which is of the order of $g_m r_o^2$ (see Eq. 13.75) is what makes it possible to realize a relatively high voltage gain in a single amplifier stage. However, such a high output resistance may be

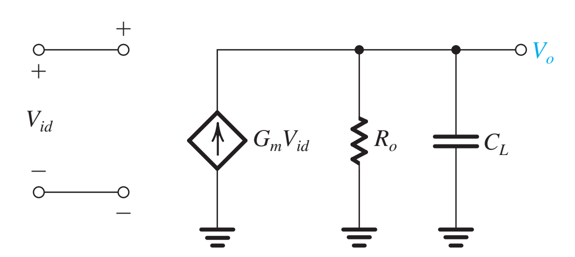


Figure 13.11 Small-signal equivalent circuit of the folded-cascode CMOS amplifier. Note that this circuit is in effect an operational transconductance amplifier (OTA).

a cause of concern to the reader; after all, in Chapter 2, we stated that an ideal op amp has a zero output resistance! To alleviate this concern somewhat, let us find the closed-loop output resistance of a unity-gain follower formed by connecting the output terminal of the circuit of Fig. 13.10 back to the negative input terminal. Since this feedback is of the voltage sampling type, it reduces the output resistance by the factor $(1 + A\beta)$, where $A = A_v$ and $\beta = 1$, that is,

$$R_{of} = \frac{R_o}{1 + A_v} \simeq \frac{R_o}{A_v} \tag{13.78}$$

Substituting for A_v from Eq. (13.76) gives

$$R_{of} \simeq \frac{1}{G_m} \tag{13.79}$$

which is a general result that applies to any OTA to which 100% voltage feedback is applied. For our particular circuit, $G_m = g_{m1}$, thus

$$R_{of} = 1/g_{m1} \tag{13.80}$$

Since g_{m1} is of the order of 1 mA/V, R_{of} will be of the order of 1 kΩ. Although this is not very small, it is reasonable in view of the simplicity of the op-amp circuit as well as the fact that this type of op amp is not usually intended to drive low-valued resistive loads.

EXERCISE

13.11 The CMOS op amp of Figs. 13.9 and 13.10 is fabricated in a process for which $V'_{An} = |V'_{Ap}| =$ 20 V/μm. If all devices have 1-μm channel length and are operated at equal overdrive voltages of 0.2-V magnitude, $I = 240$ μA, and $I_B = 150$ μA, find the voltage gain and the value of R_o obtained.
Ans. 16,000 V/V; 13.3 MΩ

13.2.4 Frequency Response

From Section 10.5, we know that one of the advantages of the cascode configuration is its excellent high-frequency response. It has three poles: one at the input, one at the connection between the CS and CG transistors (i.e., at the source terminals of Q_3 and Q_4), and the third at the output terminal. Normally, the first two poles are at very high frequencies, especially when the resistance of the signal generator that feeds the differential pair is small. Since the primary purpose of CMOS op amps is to feed capacitive loads, C_L is usually large, and the pole at the output becomes dominant. Even if C_L is not large, we can increase it deliberately to give the op amp a dominant pole and thus an open-loop gain that decreases at the uniform rate of −20 dB/decade down to the unity-gain frequency f_t, ensuring stable operation when feedback is applied. From Fig. 13.11 we can write

$$\frac{V_o}{V_{id}} = \frac{G_m R_o}{1 + sC_L R_o} \tag{13.81}$$

Thus, the dominant pole has a frequency f_P,

$$f_P = \frac{1}{2\pi C_L R_o} \tag{13.82}$$

and the unity-gain frequency f_t will be

$$f_t = G_m R_o f_P = \frac{G_m}{2\pi C_L} \tag{13.83}$$

From a design point of view, the value of C_L should be such that at $f = f_t$ the excess phase resulting from the nondominant poles is small enough to permit the required phase margin to be achieved. If C_L is not large enough to achieve this purpose, it can be augmented.

It is important to note the difference in the effect of increasing the load capacitance on the operation of each of the two op-amp circuits we have studied. In the two-stage circuit, if C_L is increased, the frequency of the second pole decreases, the excess phase shift at $f = f_t$ increases, and the phase margin is reduced. Here, on the other hand, when C_L is increased, f_t decreases, but the phase margin increases. In other words, a heavier capacitive load decreases the bandwidth of the folded-cascode amplifier but does not impair its response (which happens when the phase margin decreases). Of course, if an increase in C_L is anticipated in the two-stage op-amp case, the designer can increase C_C, thus decreasing f_t and restoring the phase margin to its required value.

13.2.5 Slew Rate

As discussed in Section 13.1.6, slewing occurs when a large differential signal appears at the op-amp input. Refer to Fig. 13.9 and consider the case when V_{id} is large and turns Q_2 off. Transistor Q_1 will then attempt to conduct the entire bias current I. This, however, would not be possible, since I is usually larger than I_B. Consideration of the drain node of Q_1 indicates that for the node equation to be satisfied, not only must the current in Q_3 reduce to zero, but also the current of Q_1 must reduce to equal I_B. For this to happen, both Q_1 and the transistor supplying I must enter the triode mode of operation, and the voltages at their drains must fall accordingly. Now, the zero current in Q_3 causes the input current of the mirror to be zero, and correspondingly its output current, in the drain of Q_6, will be zero. Meanwhile, the zero current in the drain of Q_2 forces the entire current I_B to flow through Q_4 and into C_L. This causes the output voltage v_O to ramp with a slope of I_B/C_L, which is the slew rate,

$$SR = \frac{I_B}{C_L} \tag{13.84}$$

After the slewing process is completed, before the amplifier can return to its normal linear operation, both Q_1 and the transistor that supplies the bias current I must leave the triode mode and return to the saturation mode of operation. This, however, can take some time and may introduce additional distortion in the output signal. As usual, however, creative circuit design comes to the rescue! Problem 13.27 investigates an ingenious way to deal with this issue.

Example 13.2

Consider a design of the folded-cascode op amp of Fig. 13.10 for which $I = 240\ \mu A$, $I_B = 150\ \mu A$, and $|V_{OV}|$ for all transistors is 0.25 V. Assume that the fabrication process provides $k_n' = 100\ \mu A/V^2$, $k_p' = 40\ \mu A/V^2$, $|V_A'| = 20$ V/μm, $V_{DD} = V_{SS} = 2.5$ V, and $|V_t| = 0.75$ V. Let all transistors have $L = 1$ μm and assume that $C_L = 5$ pF. Find I_D, g_m, r_o, and W/L for all transistors. Find the allowable range of V_{ICM} and of

the output voltage swing. Determine the values of A_v, f_t, f_P, and SR. What is the power dissipation of the op amp?

Solution

From the given values of I and I_B we can determine the drain current I_D for each transistor. The transconductance of each device is found using

$$g_m = \frac{2I_D}{V_{OV}} = \frac{2I_D}{0.25}$$

and the output resistance r_o from

$$r_o = \frac{|V_A|}{I_D} = \frac{20}{I_D}$$

The W/L ratio for each transistor is determined from

$$\left(\frac{W}{L}\right)_i = \frac{2I_{Di}}{k'V_{OV}^2}$$

The results are as follows:

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9	Q_{10}	Q_{11}
I_D (μA)	120	120	30	30	30	30	30	30	150	150	240
g_m (mA/V)	0.96	0.96	0.24	0.24	0.24	0.24	0.24	0.24	1.2	1.2	1.92
r_o (kΩ)	167	167	667	667	667	667	667	667	133	133	83
W/L	38.4	38.4	24	24	9.6	9.6	9.6	9.6	120	120	76.8

Note that for all transistors,

$$g_m r_o = 160\text{ V/V}$$

$$V_{GS} = 1.0\text{ V}$$

Using the expression in Eq. (13.66), the input common-mode range is found to be

$$-1.25\text{ V} \le V_{ICM} \le 3\text{ V}$$

The output voltage swing is found using Eqs. (13.68) and (13.69) to be

$$-1.25\text{ V} \le v_O \le 2\text{ V}$$

Example 13.2 *continued*

To obtain the voltage gain, we first determine R_{o4} using Eq. (13.73) as

$$R_{o4} = 160(167 \parallel 133) = 11.85\,\text{M}\Omega$$

and R_{o6} using Eq. (13.74) as

$$R_{o6} = 106.7\,\text{M}\Omega$$

The output resistance R_o can then be found as

$$R_o = R_{o4} \parallel R_{o6} = 10.7\,\text{M}\Omega$$

and the voltage gain

$$\begin{aligned} A_v &= G_m R_o = 0.96 \times 10^{-3} \times 10.7 \times 10^6 \\ &= 10,240\,\text{V/V} \end{aligned}$$

The unity-gain bandwidth is found using Eq. (13.83),

$$f_t = \frac{0.96 \times 10^{-3}}{2\pi \times 5 \times 10^{-12}} = 30.6\text{ MHz}$$

Thus, the dominant-pole frequency must be

$$f_P = \frac{f_t}{A_v} = \frac{30.6\text{ MHz}}{10,240} = 3\text{ kHz}$$

The slew rate can be determined using Eq. (13.84),

$$SR = \frac{I_B}{C_L} = \frac{150 \times 10^{-6}}{5 \times 10^{-12}} = 30\text{ V}/\mu s$$

Finally, to determine the power dissipation we note that the total current is $300\,\mu\text{A} = 0.3$ mA, and the total supply voltage is 5 V, thus

$$P_D = 5 \times 0.3 = 1.5\,\text{mW}$$

13.2.6 Increasing the Input Common-Mode Range: Rail-to-Rail Input Operation

In Section 13.2.2 we found that while the upper limit on the input common-mode range exceeds the supply voltage V_{DD}, the magnitude of the lower limit is significantly lower than V_{SS}. The opposite situation occurs if the input differential amplifier is made up of PMOS transistors. It follows that an NMOS and a PMOS differential pair placed in parallel would provide an input stage with a common-mode range that exceeds the power-supply voltage in both directions. This is known as rail-to-rail input operation. Figure 13.12 shows such an

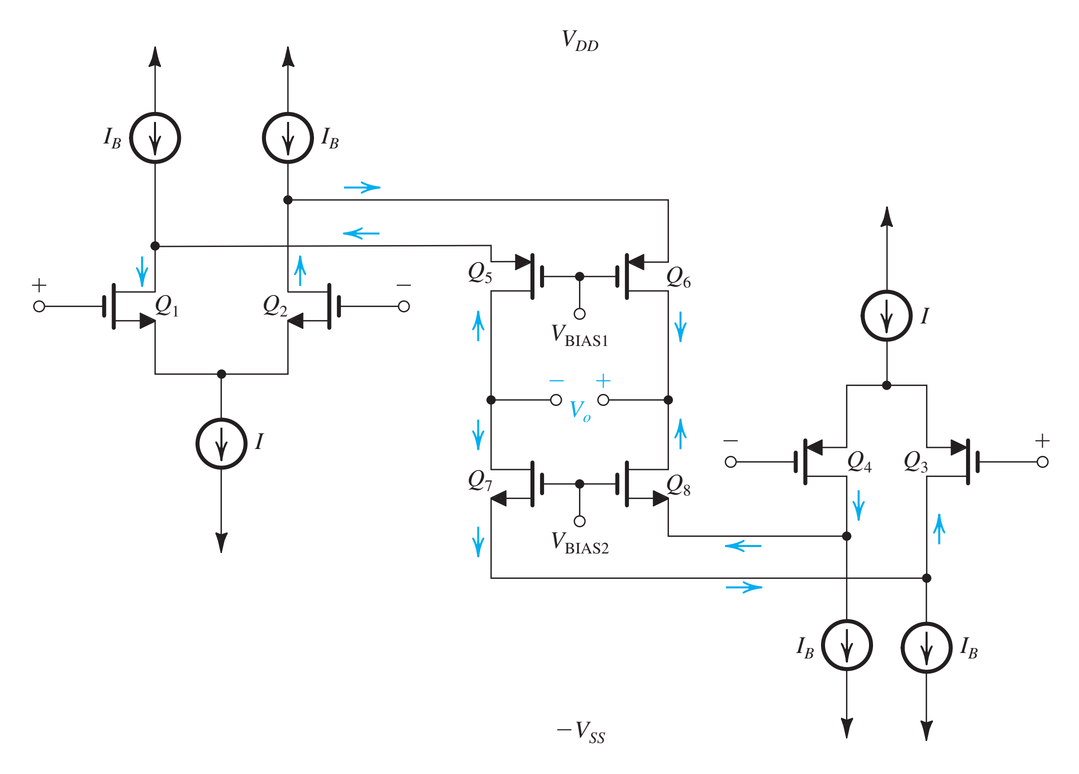


Figure 13.12 A folded-cascode op amp that employs two parallel complementary input stages to achieve rail-to-rail input common-mode operation. Note that the two "+" terminals are connected together and the two "–" terminals are connected together.

arrangement. To keep the diagram simple, we have not shown the parallel connection of the two differential pairs: The two positive-input terminals are to be connected together and the two negative-input terminals are to be tied together. Transistors Q_5 and Q_6 are the cascode transistors for the Q_1–Q_2 pair, and transistors Q_7 and Q_8 are the cascode devices for the Q_3–Q_4 pair. The output voltage V_o is shown taken differentially between the drains of the cascode devices. To obtain a single-ended output, a differential-to-single-ended conversion circuit should be connected in cascade.

Figure 13.12 indicates by arrows the direction of the current increments that result from the application of a positive differential input signal V_{id}. Each of the current increments indicated is equal to $G_m(V_{id}/2)$ where $G_m = g_{m1} = g_{m2} = g_{m3} = g_{m4}$. Thus the total current feeding each of the two output nodes will be $G_m V_{id}$. Now, if the output resistance between each of the two nodes and ground is denoted R_o, the output voltage will be

$$V_o = 2G_m R_o V_{id}$$

Thus the voltage gain will be

$$A_v = 2G_m R_o \tag{13.85}$$

This, however, assumes that both differential pairs will be operating simultaneously. This in turn occurs only over a limited range of V_{ICM}. Over the remainder of the input common-mode range, only one of the two differential pairs will be operational, and the gain drops to half of the

value in Eq. (13.85). This rail-to-rail, folded-cascode structure is utilized in a commercially available op amp.[4]

EXERCISE

13.12 For the circuit in Fig. 13.12, assume that all transistors, including those that implement the current sources, are operating at equal overdrive voltages of 0.3-V magnitude and have $|V_t| = 0.7$ V and that $V_{DD} = V_{SS} = 2.5$ V.
(a) Find the range over which the NMOS input stage operates.
(b) Find the range over which the PMOS input stage operates.
(c) Find the range over which both operate (the overlap range).
(d) Find the input common-mode range.
(Note that to operate properly, each of the current sources requires a minimum voltage of $|V_{OV}|$ across its terminals.)
Ans. –1.2 V to +2.9 V; –2.9 V to +1.2 V, –1.2 V to +1.2 V; –2.9 V to +2.9 V

13.2.7 Increasing the Output Voltage Range: The Wide-Swing Current Mirror

In Section 13.2.2 it was found that while the output voltage of the circuit of Fig. 13.10 can swing to within $2|V_{OV}|$ of V_{DD}, the cascode current mirror limits the negative swing to $[2|V_{OV}| + V_t]$ above $-V_{SS}$. In other words, the cascode mirror reduces the voltage swing by V_t volts. This point is further illustrated in Fig. 13.13(a), which shows a cascode mirror (with $V_{SS} = 0$, for simplicity) and indicates the voltages that result at the various nodes. Observe that because the voltage at the gate of Q_3 is $2V_t + 2V_{OV}$, the minimum voltage permitted at the output (while Q_3 remains saturated) is $V_t + 2V_{OV}$, hence the extra V_t. Also, observe that Q_1 is operating with a drain-to-source voltage $V_t + V_{OV}$, which is V_t volts greater than it needs to operate in saturation.

The observations above lead us to the conclusion that to permit the output voltage at the drain of Q_3 to swing as low as $2V_{OV}$, we must lower the voltage at the gate of Q_3 from $2V_t + 2V_{OV}$ to $V_t + 2V_{OV}$. This is exactly what is done in the modified mirror circuit in Fig. 13.13(b): The gate of Q_3 is now connected to a bias voltage $V_{\text{BIAS}} = V_t + 2V_{OV}$. Thus the output voltage can go down to $2V_{OV}$ with Q_3 still in saturation. Also, the voltage at the drain of Q_1 is now V_{OV} and thus Q_1 is operating at the edge of saturation. The same is true of Q_2 and thus the current tracking between Q_1 and Q_2 will be assured. Note, however, that we can no longer connect the gate of Q_2 to its drain. Rather, it is connected to the drain of Q_4. This establishes a voltage of $V_t + V_{OV}$ at the drain of Q_4, which is sufficient to operate Q_4 in saturation (as long as V_t is greater than V_{OV}, which is usually the case). This circuit is known as the **wide-swing current mirror**. Finally, note that Fig. 13.13(b) does not show the circuit for generating V_{BIAS}. There are a number of possible circuits to accomplish this task, one of which is explored in Exercise 13.13.

[4] The Texas Instruments OPA357.

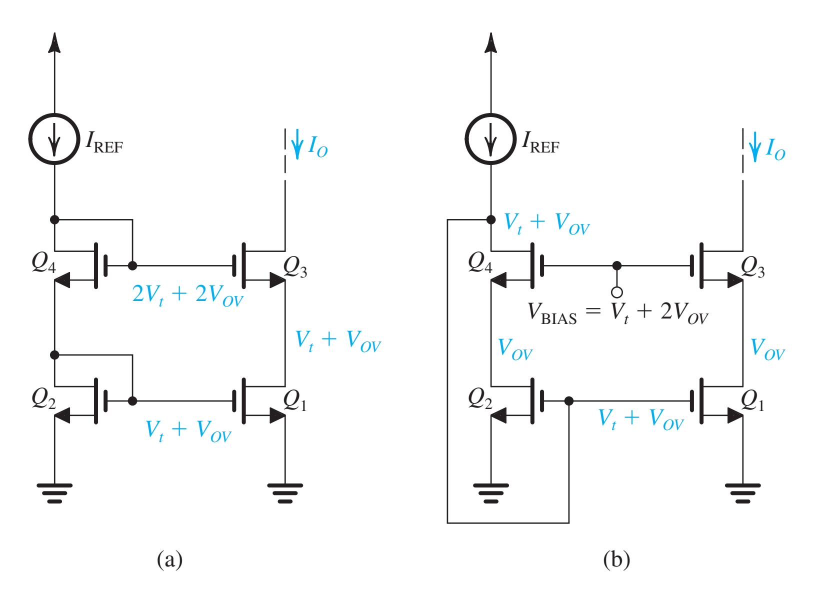


Figure 13.13 **(a)** Cascode current mirror with the voltages at all nodes indicated. Note that the minimum voltage allowed at the output is $V_t + 2V_{OV}$. **(b)** A modification of the cascode mirror that results in the reduction of the minimum output voltage to V_{OV}. This is the wide-swing current mirror. The circuit requires a bias voltage V_{BIAS}.

EXERCISE

13.13 Show that if transistor Q_5 in the circuit of Fig. E13.13 has a W/L ratio equal to one-quarter that of the transistors in the wide-swing current mirror of Fig. 13.13(b), and provided the same value of I_{REF} is utilized in both circuits, then the voltage generated, V_5, is $V_t + 2V_{OV}$, which is the value of V_{BIAS} needed for the gates of Q_3 and Q_4.

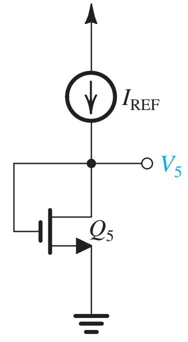


Figure E13.13

13.3 The 741 BJT Op Amp

Our study of BJT op amps is in two parts: The first part, in this section, is focused on an analog IC classic, the 741 op-amp circuit; the second part, in Section 13.4, presents some of the more recent design techniques.

13.3.1 The 741 Circuit

Figure 13.14 shows the 741 op-amp circuit. In keeping with the IC design philosophy, the circuit uses a large number of transistors, but relatively few resistors and only one capacitor. This philosophy is dictated by the economics (silicon area, ease of fabrication, quality of realizable components) of the fabrication of active and passive components in IC form (see Section 8.1 and Appendix A).

As in the case of most general-purpose IC op amps, the 741 requires two power supplies, $+V_{CC}$ and $-V_{EE}$. Normally, $V_{CC} = V_{EE} = 15$ V, but the circuit also operates satisfactorily with the power supplies reduced to much lower values (such as ± 5 V).[5] It is important to observe that no circuit node is connected to ground, the common terminal of the two supplies.

With a relatively large circuit like that shown in Fig. 13.14, the first step in the analysis is to identify its recognizable parts and their functions. Thus, we begin with a qualitative description of the circuit. Our description is aided by the division of the circuit into its various parts, as indicated in the diagram.

Bias Circuit The reference bias current of the 741 circuit, I_{REF}, is generated in the branch at the extreme left of Fig. 13.14, consisting of the two diode-connected transistors Q_{11} and Q_{12} and the resistance R_5. Using a Widlar current source formed by Q_{11}, Q_{10}, and R_4, bias current for the first stage is generated in the collector of Q_{10}. Another current mirror formed by Q_8 and Q_9 takes part in biasing the first stage.

The reference bias current I_{REF} is used to provide two proportional currents in the collectors of Q_{13}. This double-collector lateral[6] *pnp* transistor can be thought of as two transistors whose base–emitter junctions are connected in parallel. Thus Q_{12} and Q_{13} form a two-output current mirror: One output, the collector of Q_{13B}, provides bias current and acts as a current-source load for Q_{17}, and the other output, the collector of Q_{13A}, provides bias current for the output stage of the op amp.

Two more transistors, Q_{18} and Q_{19}, take part in the dc bias process. The purpose of Q_{18} and Q_{19} is to establish two V_{BE} drops between the bases of the output transistors Q_{14} and Q_{20}.

Short-Circuit-Protection Circuitry The 741 circuit includes a number of transistors that are normally off and conduct only if one attempts to draw a large current from the op-amp output terminal. This happens, for example, if the output terminal is short-circuited to one of the two supplies. The short-circuit-protection network (shown in color in Fig. 13.14) consists of R_6, R_7, Q_{15}, Q_{21}, Q_{24}, R_{11}, and Q_{22}. In the following we shall assume that these transistors are off. Operation of the short-circuit-protection network will be explained in Section 13.3.3.

[5]The 741 is fabricated in what is known as the “standard high-voltage technology.” See Appendix K for the parameter values of devices fabricated in this process.

[6]See Appendix A for a description of lateral *pnp* transistors. Also, their characteristics are given in Appendix K.

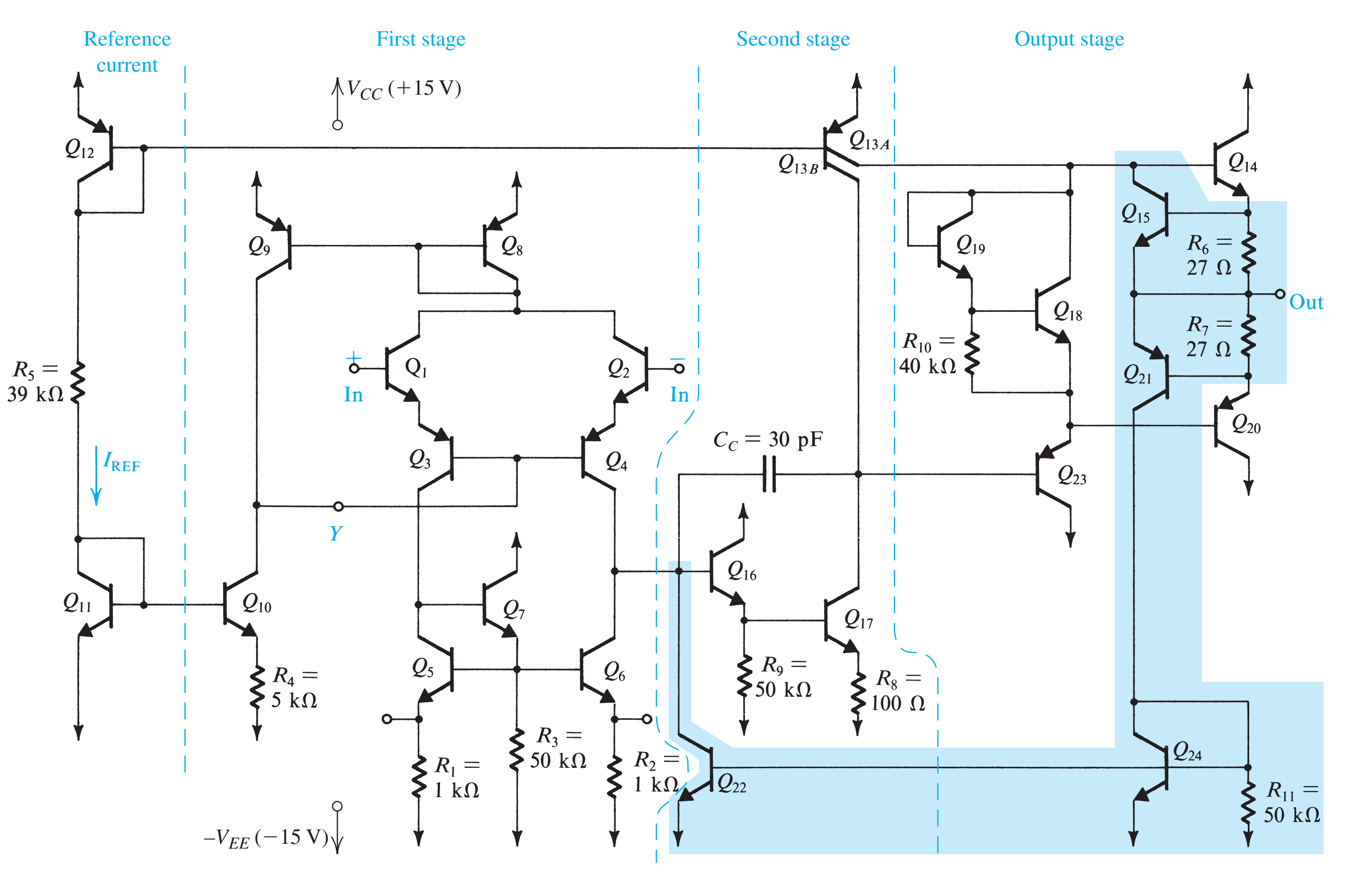


Figure 13.14 The 741 op-amp circuit: Q_{11}, Q_{12}, and R_5 generate a reference bias current, I_{REF}. Q_{10}, Q_9, and Q_8 bias the input stage, which is composed of Q_1 to Q_7. The second gain stage is composed of Q_{16} and Q_{17} with Q_{13B} acting as active load. The class AB output stage is formed by Q_{14} and Q_{20} with biasing devices Q_{13A}, Q_{18}, and Q_{19}, and an input buffer Q_{23}. Transistors Q_{15}, Q_{21}, Q_{24}, and Q_{22} serve to protect the amplifier against output short circuits and are normally cut off.

The Input Stage The 741 circuit consists of three stages: an input differential stage, an intermediate single-ended high-gain stage, and an output-buffering stage. The input stage consists of transistors Q_1 through Q_7, with biasing performed by Q_8, Q_9, and Q_{10}. Transistors Q_1 and Q_2 act as emitter followers, causing the input resistance to be high and delivering the differential input signal to the differential common-base amplifier formed by Q_3 and Q_4. Thus the input stage is the differential version of the common-collector, common-base configuration discussed in Section 8.7.3.

Transistors Q_5, Q_6, and Q_7 and resistors R_1, R_2, and R_3 form the load circuit of the input stage. This is an elaborate current-mirror-load circuit, which we will analyze in Section 13.3.3. The circuit is based on the base-current-compensated mirror studied in Section 8.2.3, but it includes two emitter-degeneration resistors R_1 and R_2, and a large resistor R_3 in the emitter of Q_7. As is the case with current-mirror loads, this circuit not only provides a high-resistance load for Q_4 but also converts the signal from differential to single-ended form with no loss in gain or common-mode rejection. The output of the input stage is taken single-endedly at the collector of Q_4.

As mentioned in Section 9.6.2, every op-amp circuit includes a *level shifter* whose function is to shift the dc level of the signal so that the signal at the op-amp output can swing positive and negative. In the 741, level shifting is done in the first stage using the lateral *pnp* transistors Q_3 and Q_4. Although lateral *pnp* transistors have poor high-frequency performance, their use in the common-base configuration (which is known to have good high-frequency response) does not seriously impair the op-amp frequency response.

The use of the lateral *pnp* transistors Q_3 and Q_4 in the first stage results in an added advantage: protection of the input-stage transistors Q_1 and Q_2 against emitter–base junction breakdown. Since the emitter–base junction of an *npn* transistor breaks down at about 7 V of reverse bias (see Section 6.4.1), regular *npn* differential stages suffer such a breakdown if, say, the supply voltage is accidentally connected between the input terminals. Lateral *pnp* transistors, however, have high emitter–base breakdown voltages (about 50 V), and because they are connected in series with Q_1 and Q_2, they provide protection of the 741 input transistors, Q_1 and Q_2.

Finally, note that except for using input buffer transistors, the 741 input stage is essentially a current-mirror-loaded differential amplifier. It is quite similar to the input stage of the CMOS amplifier in Fig. 13.1.

The Second Stage The second or intermediate stage is composed of Q_{16}, Q_{17}, Q_{13B}, and the two resistors R_8 and R_9. Transistor Q_{16} acts as an emitter follower, thus giving the second stage a high input resistance. This minimizes the loading on the input stage and avoids loss of gain. Also, adding Q_{16} with its 50-kΩ emitter resistance (which is similar to Q_7 and R_3) increases the symmetry of the first stage and thus improves its CMRR. Transistor Q_{17} acts as a common-emitter amplifier with a 100-Ω resistor in the emitter. Its load is composed of the high output resistance of the *pnp* current source Q_{13B} in parallel with the input resistance of the output stage (seen looking into the base of Q_{23}). Using a transistor current source as a load resistance (*active load*) enables one to obtain high gain without resorting to the use of large resistances, which would occupy a large chip area and require large power-supply voltages.

The output of the second stage is taken at the collector of Q_{17}. Capacitor C_C is connected in the feedback path of the second stage to provide frequency compensation using the Miller compensation technique studied in Section 11.10.3. It will be shown in Section 13.3.4 that the relatively small capacitor C_C gives the 741 a dominant pole at about 4 Hz. Furthermore, pole splitting causes other poles to be shifted to much higher frequencies, giving the op amp a uniform –20-dB/decade gain rolloff with a unity-gain bandwidth of about 1 MHz. It should

be pointed out that although C_C is small in value, the chip area that it occupies is about 13 times that of a standard *npn* transistor!

THE CREATOR OF THE μA741—DAVID FULLAGAR:

David Fullagar was at Fairchild Semiconductor in 1967 when he designed the μA741, perhaps the most successful op amp ever. Fairchild, TI, and National still sell updated versions of this ubiquitous device. Fullagar, educated at Cambridge, U.K., and formerly employed at Ferranti, had joined Fairchild in 1966 following Widlar's departure after the μA702 and μA709 designs. Fullagar's μA741 creation incorporated internal compensation, short-circuit protection, and a novel high-impedance input stage to resolve shortcomings of the earlier designs. After leaving Fairchild, he joined Intersil as the company's first analog IC designer. The engineer-designer cofounded and became a vital technical contributor to Maxim Integrated Products in 1983; he retired in 1999.

The Output Stage The purpose of the output stage (Chapter 12) is to provide the amplifier with a low output resistance. In addition, the output stage should be able to supply relatively large load currents without dissipating an unduly large amount of power in the IC. The 741 uses an efficient class AB output stage, which we shall study in Section 13.3.3.

The output stage of the 741 consists of the complementary pair Q_{14} and Q_{20}, where Q_{20} is a *substrate pnp* (see Appendix A). Transistors Q_{18} and Q_{19} are fed by current source Q_{13A} and bias the output transistors Q_{14} and Q_{20}. Transistor Q_{23} (which is another substrate *pnp*) acts as an emitter follower, thus minimizing the loading effect of the output stage on the second stage.

Device Parameters In the following sections and in the exercises and end-of-chapter problems we shall carry out a detailed analysis of the 741 circuit. For the standard *npn* and *pnp* transistors, the following parameters will be used:

$$npn:\quad I_S = 10^{-14}\text{A},\ \beta = 200,\ V_A = 125\text{ V}$$
$$pnp:\quad I_S = 10^{-14}\text{A},\ \beta = 50,\ V_A = 50\text{ V}$$

In the 741 circuit the nonstandard devices are Q_{13}, Q_{14}, and Q_{20}. Transistor Q_{13} will be assumed to be equivalent to two transistors, Q_{13A} and Q_{13B}, with parallel base–emitter junctions and having the following saturation currents:

$$I_{SA} = 0.25 \times 10^{-14}\text{ A}\quad I_{SB} = 0.75 \times 10^{-14}\text{ A}$$

Transistors Q_{14} and Q_{20} will be assumed to each have an area three times that of a standard device. Output transistors usually have relatively large areas, to be able to supply large load currents and dissipate relatively large amounts of power with only a moderate increase in device temperature.

EXERCISE

13.14 For the standard *npn* transistor whose parameters are given in Section 13.3.1, find approximate values for the following parameters at $I_C = 0.1$ mA: V_{BE}, g_m, r_e, r_π, and r_o.
Ans. 575 mV; 4 mA/V; 250 Ω; 50 kΩ; 1.25 MΩ

13.3.2 DC Analysis

In this section, we shall carry out a dc analysis of the 741 circuit to determine the bias point of each device. For the dc analysis of an op-amp circuit, the input terminals are grounded. Theoretically speaking, this should result in zero dc voltage at the output. However, because the op amp has very large gain, any slight approximation in the analysis will show that the output voltage is far from being zero and is close to either $+V_{CC}$ or $-V_{EE}$. In actual practice, an op amp left open-loop will have an output voltage saturated close to one of the two supplies. To overcome this problem in the dc analysis, it will be assumed that the op amp is connected in a negative feedback loop that stabilizes the output dc voltage to zero volts.

Reference Bias Current The reference bias current I_{REF} is generated in the branch composed of the two diode-connected transistors Q_{11} and Q_{12} and resistor R_5. Thus,

$$I_{\text{REF}} = \frac{V_{CC} - V_{EB12} - V_{BE11} - (-V_{EE})}{R_5}$$

For $V_{CC} = V_{EE} = 15$ V and $V_{BE11} = V_{EB12} \simeq 0.7$ V, we have $I_{\text{REF}} = 0.73$ mA.

Input-Stage Bias Transistors Q_{11} and Q_{10} and resistor R_4 form a Widlar current source (Section 8.6.4), thus

$$V_T \ln \frac{I_{\text{REF}}}{I_{C10}} = I_{C10} R_4 \tag{13.86}$$

EXERCISE

13.15 Use Eq. (13.86) to determine the value of I_{C10} by trial and error. Note that $I_{\text{REF}} = 0.73$ mA and $R_4 = 5$ kΩ.
Ans. $I_{C10} = 19\ \mu$A

Having determined I_{C10}, we proceed to determine the dc current in each of the input-stage transistors. For this purpose, we show in Fig. 13.15 the centerpiece of the input stage: As will

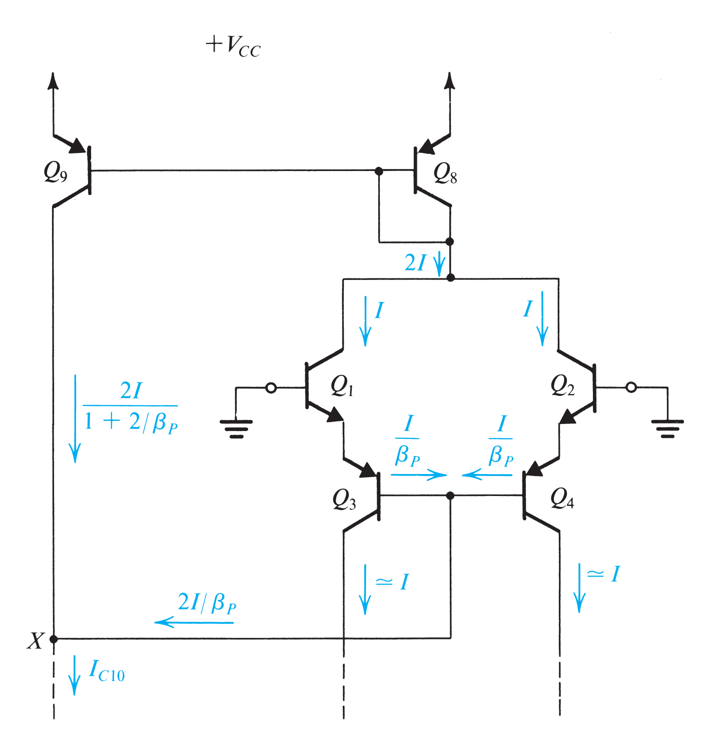


Figure 13.15 The dc analysis of the 741 input stage.

be seen shortly, this is a negative-feedback circuit that stabilizes the bias current of each of Q_1 to Q_4 at a value approximately equal to $I_{C10}/2$. Refer to the analysis indicated in the diagram (where β_N is assumed to be high). The sum of the collector currents of Q_1 and Q_2 ($2I$) is fed to (or sensed by) the input of the current mirror Q_8–Q_9. The output current of the mirror, which for large β_P is approximately equal to $2I$, is compared to I_{C10} at node X. The difference between the two currents ($2I/\beta_P$) establishes the base currents of Q_3 and Q_4. This is the error signal of the feedback loop. For large β_P, this current approaches zero and a node equation at X gives $2I \simeq I_{C10}$, and thus $I \simeq I_{C10}/2$.

To verify the action of the negative-feedback loop in stabilizing the value of I, assume that for some reason I increases. We see that the input current of the Q_8–Q_9 mirror increases and, correspondingly, its output current increases. Assuming that I_{C10} remains constant, consideration of node X reveals that the base currents in Q_3 and Q_4 must decrease. This in turn decreases the value of I, which is opposite to the originally assumed change.

EXERCISES

13.16 Using the value of I_{C10} found in Exercise 13.15, find the value of the bias current of each of Q_1, Q_2, Q_3, and Q_4.
Ans. 9.5 μA

13.17 It is required to determine the loop gain of the feedback loop in Fig. 13.15. Break the loop at the input of the Q_8–Q_9 mirror. Since the input resistance of the mirror is low, ground the connection of the collectors of Q_1 and Q_2. Apply an input test current i_t to the current mirror and find the feedback current that appears in the combined connection of the collectors of Q_1 and Q_2. Assume I_{C10} remains constant.
Ans. Loop gain $\simeq \beta_P$

Continuing with the dc analysis of the input stage, we show in Fig. 13.16 the current-mirror load (Q_5, Q_6, and Q_7) and the input transistor of the second stage (Q_{16}). The current-mirror load is fed by $I_{C3} = I_{C4} \simeq I$. The analysis is illustrated in the figure and shows that for large β_N, each of Q_5 and Q_6 is biased at a current approximately equal to I. The bias current of Q_7 is somewhat higher, as shown in Exercise 13.18.

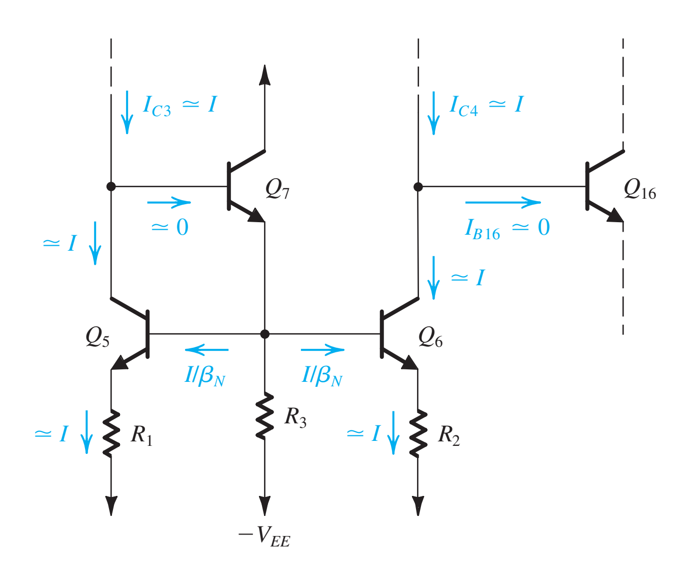


Figure 13.16 Continuation of the dc analysis of the 741 input stage.

EXERCISES

13.18 Refer to Fig. 13.16 and recall that $I = 9.5\ \mu$A, $R_1 = R_2 = 1\ \text{k}\Omega$, $R_3 = 50\ \text{k}\Omega$, $\beta_N = 200$, and I_S (for all three transistors) is 10^{-14} A. Find V_{BE6}, V_{B6}, and I_{C7}.
Ans. 517 mV; 526.5 mV; 10.5 μA

13.19 Recalling from Chapters 2 and 9 that the input bias current of an op amp is the average of its two input currents, thus

$$I_B = \frac{1}{2}(I_{B1} + I_{B2})$$

and the input offset current is

$$I_{OS} = |I_{B1} - I_{B2}|$$

find I_B and I_{OS} for the 741 if β_1 and β_2 are nominally 200 but can deviate from nominal by as much as $\pm 5\%$.

Ans. 47.5 nA; 4.75 nA

Input Common-Mode Range The **input common-mode range** is the range of input common-mode voltages over which the input stage remains in the linear active mode. Refer to Fig. 13.14. We see that in the 741 circuit the input common-mode range is determined at the upper end by saturation of Q_1 and Q_2, and at the lower end by saturation of Q_3 and Q_4.

EXERCISE

13.20 Neglect the voltage drops across R_1 and R_2 and assume that $V_{CC} = V_{EE} = 15$ V. Show that the input common-mode range of the 741 is approximately -12.9 V to $+14.7$ V. (Assume that $V_{BE} \simeq 0.6$ V and that to avoid saturation $V_{CB} \geq -0.3$ V for an *npn* transistor, and $V_{BC} \geq -0.3$ V for a *pnp* transistor.)

Second-Stage Bias Reference to Fig. 13.14 shows that if we neglect the base current of Q_{23}, the collector current of Q_{17} will be equal to the current supplied by Q_{13B}. We can then use I_{C17} to determine V_{BE17}, V_{B17}, the current through R_9 and hence I_{E16}, and finally $I_{C16} \simeq I_{E16}$.

EXERCISE

13.21 Recalling that Q_{13B} has a scale current 0.75 times that of Q_{12}, find I_{C13B} and hence I_{C17}. Assume $\beta_P \gg 1$. Then determine V_{BE17}, I_{C16}, and I_{B16}. (Recall that $I_{\text{REF}} = 0.73$ mA, $I_S = 10^{-14}$ A, and $\beta_N = 200$.)

Ans. 550 μA; 550 μA; 618 mV; 16.2 μA; 0.08 μA

Output-Stage Bias Figure 13.17 shows the output stage of the 741 with the short-circuit-protection circuitry omitted. Current source Q_{13A} delivers a current of $0.25I_{\text{REF}}$ (because I_S of Q_{13A} is 0.25 times the I_S of Q_{12}) to the network composed of Q_{18}, Q_{19}, and R_{10}. As mentioned in Section 13.3.1, the purpose of the Q_{18}–Q_{19} network is to establish two V_{BE} drops between

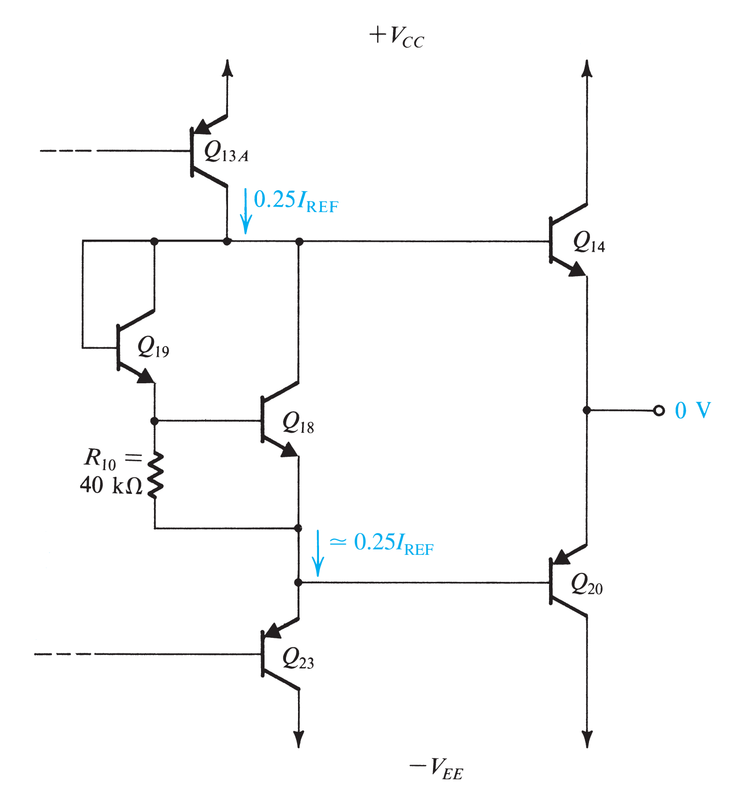


Figure 13.17 The 741 output stage without the short-circuit-protection devices.

the bases of the output transistors Q_{14} and Q_{20}. If we neglect the base currents of Q_{14} and Q_{20}, then the emitter current of Q_{23} will also be equal to $0.25I_{REF}$.

The determination of the bias currents of the output-stage transistors is illustrated by the following example.

Example 13.3

Determine I_{C23}, I_{B23}, $V_{BB} = V_{BE18} + V_{BE19}$, I_{C14}, and I_{C20}. Recall that Q_{14} and Q_{20} are nonstandard devices with $I_{S14} = I_{S20} = 3 \times 10^{-14}$ A.

Solution

Reference to Fig. 13.7 shows that

$$I_{C23} \simeq I_{E23} \simeq 0.25I_{REF} = 180\,\mu\text{A}$$

Thus we see that the base current of Q_{23} is only $180/50 = 3.6\,\mu$A, which is negligible compared to I_{C17}, as we assumed before.

If we assume that V_{BE18} is approximately 0.6 V, we can determine the current in R_{10} as 15 μA. The emitter current of Q_{18} is therefore

$$I_{E18} = 180 - 15 = 165\,\mu\text{A}$$

Also,

$$I_{C18} \simeq I_{E18} = 165\,\mu\text{A}$$

At this value of current we find that $V_{BE18} = 588$ mV, which is quite close to the value assumed. The base current of Q_{18} is $165/200 = 0.8$ μA, which can be added to the current in R_{10} to determine the Q_{19} current as

$$I_{C19} \simeq I_{E19} = 15.8\,\mu\text{A}$$

The voltage drop across the base–emitter junction of Q_{19} can now be determined as

$$V_{BE19} = V_T \ln \frac{I_{C19}}{I_S} = 530\,\text{mV}$$

The voltage drop V_{BB} can now be calculated as

$$V_{BB} = V_{BE18} + V_{BE19} = 588 + 530 = 1.118\text{ V}$$

Since V_{BB} appears across the series combination of the base–emitter junctions of Q_{14} and Q_{20}, we can write

$$V_{BB} = V_T \ln \frac{I_{C14}}{I_{S14}} + V_T \ln \frac{I_{C20}}{I_{S20}}$$

Using the calculated value of V_{BB} and substituting $I_{S14} = I_{S20} = 3 \times 10^{-14}$ A, we determine the collector currents as

$$I_{C14} = I_{C20} = 154\,\mu\text{A}$$

This is the small current (relative to the load currents that the output stage is called upon to supply) at which the class AB output stage is biased.

Summary For future reference, Table 13.1 provides a listing of the values of the collector bias currents of the 741 transistors.

Table 13.1 DC Collector Currents of the 741 Circuit (μA)

Q_1	9.5	Q_8	19	Q_{13B}	550	Q_{19}	15.8
Q_2	9.5	Q_9	19	Q_{14}	154	Q_{20}	154
Q_3	9.5	Q_{10}	19	Q_{15}	0	Q_{21}	0
Q_4	9.5	Q_{11}	730	Q_{16}	16.2	Q_{22}	0
Q_5	9.5	Q_{12}	730	Q_{17}	550	Q_{23}	180
Q_6	9.5	Q_{13A}	180	Q_{18}	165	Q_{24}	0
Q_7	10.5						

EXERCISE

13.22 If in the circuit of Fig. 13.17 the Q_{18}–Q_{19} network is replaced by two diode-connected transistors, find the current in Q_{14} and Q_{20}. Assume that the diode-connected transistors utilize standard devices with $I_S = 10^{-14}$ A, while the nonstandard Q_{14} and Q_{20} have $I_S = 3 \times 10^{-14}$ A.
Ans. 540 μA

13.3.3 Small-Signal Analysis

The Input Stage Figure 13.18 shows part of the 741 input stage for the purpose of performing small-signal analysis. Note that since the collectors of Q_1 and Q_2 are connected to a constant dc voltage, they are shown grounded. Also, the constant-current biasing of the bases of Q_3 and Q_4 is equivalent to having the common-base terminal open-circuited.

The differential signal v_i applied between the input terminals effectively appears across four equal emitter resistances connected in series—those of Q_1, Q_2, Q_3, and Q_4. As a result,

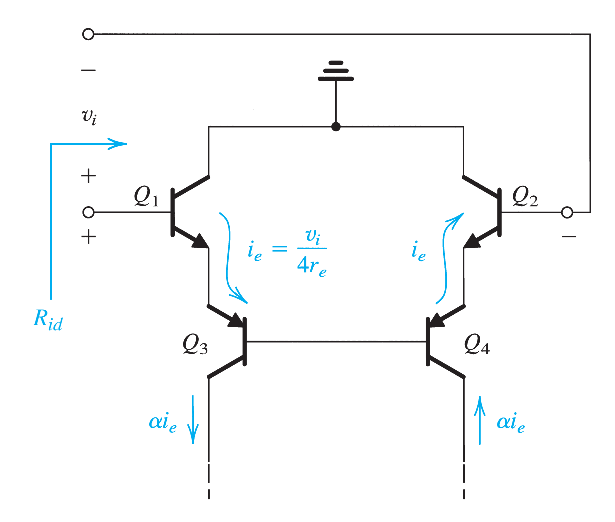


Figure 13.18 Small-signal analysis of the 741 input stage.

emitter signal currents flow as indicated in Fig. 13.18 with

$$i_e = \frac{v_i}{4r_e} \tag{13.87}$$

where r_e denotes the emitter resistance of each of Q_1 through Q_4. Thus

$$r_e = \frac{V_T}{I}$$

Thus the four transistors Q_1 through Q_4 supply the load circuit with a pair of complementary current signals αi_e, as indicated in Fig. 13.18.

The input differential resistance of the op amp can be obtained from Fig. 13.18 as

$$R_{id} = 4(\beta_N + 1)r_e \tag{13.88}$$

Proceeding with the input-stage analysis, we show in Fig. 13.19 the current-mirror-load circuit fed with the complementary pair of current signals found earlier. The analysis, together with the order of the steps in which it is performed, is indicated on the diagram. As expected, the current mirror provides an output current i_o,

$$i_o = 2\alpha i_e \tag{13.89}$$

Combining Eqs. (13.87) and (13.89) provides the transconductance of the input stage as

$$G_{m1} \equiv \frac{i_o}{v_i} = \frac{\alpha}{2r_e} = \frac{1}{2}g_{m1} \tag{13.90}$$

where g_{m1} is the transconductance of each of the four transistors Q_1, Q_2, Q_3, and Q_4.

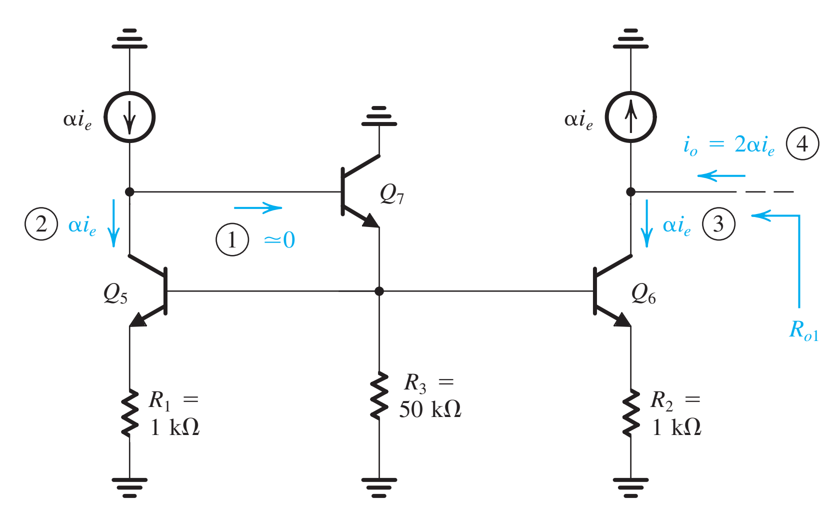


Figure 13.19 The current-mirror-load circuit of the input stage fed by the two complementary current signals generated by Q_1 through Q_4 in Fig. 13.18. Circled numbers indicate the order of the analysis steps.

EXERCISES

13.23 Recalling that each of the input-stage transistors is biased at a current $I = 9.5\ \mu A$ and that $\beta_N = 200$, find r_e, g_{m1}, G_{m1}, and R_{id}.
Ans. 2.63 kΩ; 0.38 mA/V; 0.19 mA/V; 21 MΩ

13.24 For the circuit in Fig. 13.19, find the following in terms of i_e: (a) the signal voltage at the base of Q_6; (b) the signal current in the emitter of Q_7; (c) the signal current in the base of Q_7; (d) the signal voltage at the base of Q_7; (e) the input resistance seen by the left-hand-side signal current source αi_e. For simplicity, assume that $I_{C7} \simeq I_{C5} = I_{C6}$, and use the results of Exercise 13.23.
Ans. (a) 3.63 kΩ × i_e; (b) $0.08 i_e$; (c) $0.0004 i_e$; (d) 3.84 kΩ × i_e; (e) 3.84 kΩ

To complete our modeling of the 741 input stage, we must find its output resistance R_{o1}. This is the resistance seen "looking back" into the output terminal of the circuit in Fig. 13.19. Thus R_{o1} is the parallel equivalent of the output resistance of the current source supplying the signal current αi_e, and the output resistance of Q_6. The first component is the resistance looking into the collector of Q_4 in Fig. 13.18. Finding this resistance is considerably simplified if we assume that the common bases of Q_3 and Q_4 are at a *virtual ground*. This of course happens only when the input signal v_i is applied in a complementary fashion. Nevertheless, making this assumption does not result in a large error.

Assuming that the base of Q_4 is at virtual ground, the resistance we are after is R_{o4}, indicated in Fig. 13.20(a). This is the output resistance of a common-base transistor that has a resistance (r_e of Q_2) in its emitter. To find R_{o4} we use the following expression (Eq. 8.70):

$$R_o = r_o\left[1 + g_m\left(R_e \| r_\pi\right)\right] \tag{13.91}$$

where $R_e = r_e$ and $r_o = V_{Ap}/I$.

The second component of the output resistance is that seen looking into the collector of Q_6 in Fig. 13.19 with the αi_e generator set to 0. Although the base of Q_6 is not at signal ground, we shall assume that the signal voltage at the base is small enough to make this approximation valid. The circuit then takes the form shown in Fig. 13.20(b), and R_{o6} can be determined using Eq. (13.91) with $R_e = R_2$.

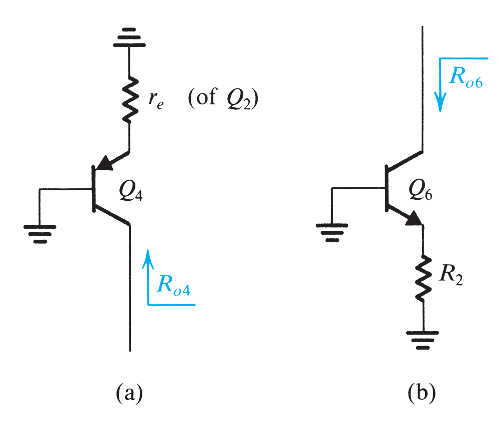


Figure 13.20 Simplified circuits for finding the two components of the output resistance R_{o1} of the first stage.

Figure 13.21 shows the equivalent circuit that we have derived for the input stage.

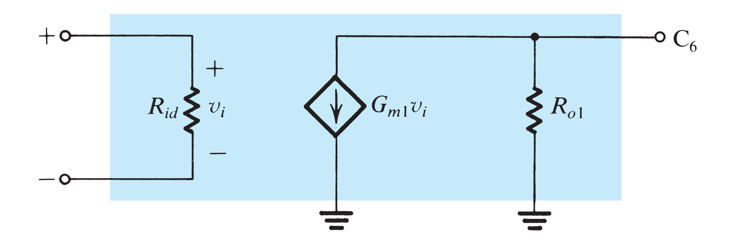


Figure 13.21 Small-signal equivalent circuit for the input stage of the 741 op amp.

EXERCISES

13.25 Find the values of R_{o4} and R_{o6} and thus the output resistance of the first stage, R_{o1}. Recall that $I = 9.5\ \mu A$, $V_{An} = 125$ V, $V_{Ap} = 50$ V, $R_2 = 1\ k\Omega$, $\beta_N = 200$, and $\beta_P = 50$.
Ans. 10.5 MΩ; 18.2 MΩ; 6.7 MΩ

13.26 Use the equivalent circuit of Fig. 13.21 together with the value of G_{m1} found in Exercise 13.23 and the value of R_{o1} found in Exercise 13.25 to determine the open-circuit voltage gain of the 741 input stage.
Ans. $|A_{vo}| = G_{m1}R_{o1} = 1273$ V/V

Example 13.4

We wish to find the input offset voltage resulting from a 2% mismatch between the resistances R_1 and R_2 in Fig. 13.14.

Solution

Consider first the situation when both input terminals are grounded, and assume that $R_1 = R$ and $R_2 = R + \Delta R$, where $\Delta R/R = 0.02$. From Fig. 13.22 we see that while Q_5 still conducts a current equal to I, the current in Q_6 will be smaller by ΔI. The value of ΔI can be found from

$$V_{BE5} + IR = V_{BE6} + (I - \Delta I)(R + \Delta R)$$

Thus

$$V_{BE5} - V_{BE6} = I\Delta R - \Delta I(R + \Delta R) \qquad (13.92)$$

The quantity on the left-hand side is in effect the change in V_{BE} due to a change in I_E of ΔI. We may therefore write

$$V_{BE5} - V_{BE6} \simeq \Delta I r_e \qquad (13.93)$$

Example 13.4 *continued*

Equations (13.92) and (13.93) can be combined to obtain

$$\frac{\Delta I}{I} = \frac{\Delta R}{R + \Delta R + r_e} \tag{13.94}$$

Substituting $R = 1\ \text{k}\Omega$ and $r_e = 2.63\ \text{k}\Omega$ shows that a 2% mismatch between R_1 and R_2 gives rise to an output current $\Delta I = 5.5 \times 10^{-3} I$. To reduce this output current to zero we have to apply an input voltage V_{OS} given by

$$V_{OS} = \frac{\Delta I}{G_{m1}} = \frac{5.5 \times 10^{-3} I}{G_{m1}} \tag{13.95}$$

Substituting $I = 9.5\ \mu\text{A}$ and $G_{m1} = 0.19\ \text{mA/V}$ results in the offset voltage $V_{OS} \simeq 0.3\ \text{mV}$.

It should be pointed out that the offset voltage calculated is only one component of the input offset voltage of the 741. Other components arise because of mismatches in transistor characteristics. The 741 offset voltage is specified to be typically 2 mV.

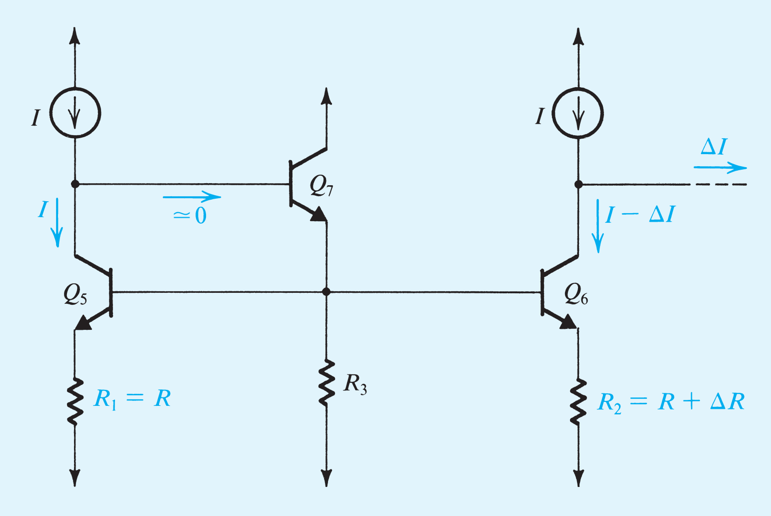


Figure 13.22 Input stage with both inputs grounded and a mismatch ΔR between R_1 and R_2.

Example 13.5

It is required to find the CMRR of the 741 input stage. Assume that the circuit is balanced except for mismatches in the current-mirror load that result in an error ϵ_m in the mirror's current-transfer ratio; that is, the ratio becomes $(1 - \epsilon_m)$.

Solution

In Section 9.5.5 we analyzed the common-mode operation of the current-mirror-loaded differential amplifier and derived an expression for its CMRR. The situation in the 741 input stage, however, differs substantially because of the feedback loop that regulates the bias current. Since this feedback loop is sensitive to the common-mode signal, as will be seen shortly, the loop operates to reduce the common-mode gain and, correspondingly, to increase the CMRR. Hence, its action is referred to as **common-mode feedback**.

Figure 13.23 shows the 741 input stage with a common-mode signal v_{icm} applied to both input terminals. We have assumed that as a result of v_{icm}, a signal current i flows as shown. Since the stage is balanced, both sides carry the same current i.

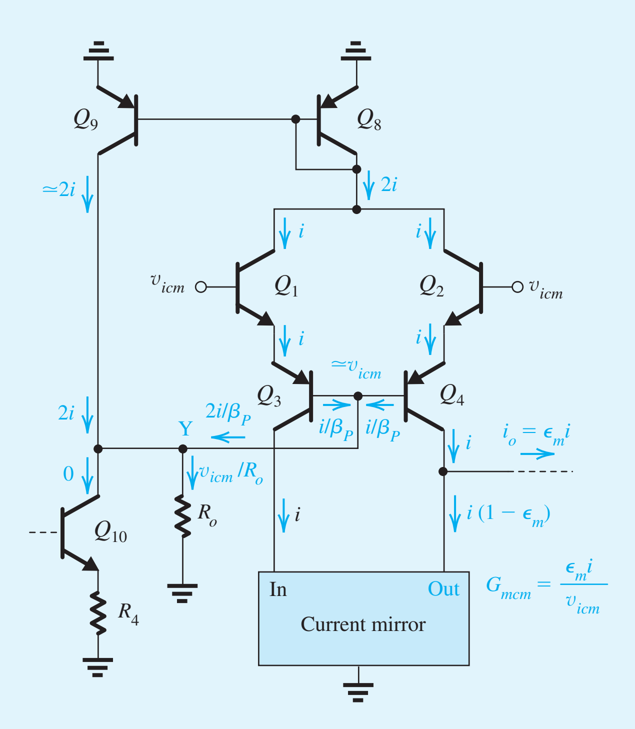


Figure 13.23 Example 13.5: Analysis of the common-mode gain of the 741 input stage. Note that $R_o = R_{o9} \| R_{o10}$ has been "pulled out" and shown separately, leaving behind ideal current sources Q_9 and Q_{10}.

Our objective now is to determine how i relates to v_{icm}. Toward that end, observe that for common-mode inputs, both sides of the differential amplifier, that is, Q_1-Q_3 and Q_2-Q_4, act as followers, delivering a signal almost equal to v_{icm} to the common-base node of Q_3 and Q_4. Now, this node Y is connected to the collectors of two current sources, Q_9 and Q_{10}. Denoting the total resistance between node Y and ground R_o, we write

$$R_o = R_{o9} \| R_{o10} \tag{13.96}$$

Example 13.5 *continued*

In Fig. 13.23 we have "pulled R_o out," thus leaving behind ideal current sources Q_9 and Q_{10}. Since the current in Q_{10} is constant, we show Q_{10} in Fig. 13.23 as having a zero incremental current. Transistor Q_9, on the other hand, provides a current approximately equal to that fed into Q_8, which is $2i$. This is the feedback current. Since Q_8 senses the *sum* of the currents in the two sides of the differential amplifier, the feedback loop operates only on the common-mode signal and is insensitive to any difference signal.

Proceeding with the analysis, we now can write a node equation at Y,

$$2i + \frac{2i}{\beta_P} = \frac{v_{icm}}{R_o} \tag{13.97}$$

Assuming $\beta_P \gg 1$, this equation simplifies to

$$i \simeq \frac{v_{icm}}{2R_o} \tag{13.98}$$

Having determined i, we now proceed to complete our analysis by finding the output current i_o. From the circuit in Fig. 13.23, we see that

$$i_o = \epsilon_m i \tag{13.99}$$

Thus the common-mode transconductance of the input stage is given by

$$G_{mcm} \equiv \frac{i_o}{v_{icm}} = \frac{\epsilon_m i}{v_{icm}}$$

Substituting for i from Eq. (13.98) gives

$$G_{mcm} = \frac{\epsilon_m}{2R_o} \tag{13.100}$$

Finally, the CMRR can be found as the ratio of the differential transconductance G_{m1} found in Eq. (13.90) and the common-mode transconductance G_{mcm},

$$\text{CMRR} \equiv \frac{G_{m1}}{G_{mcm}} = g_{m1}R_o/\epsilon_m \tag{13.101}$$

where g_{m1} is the transconductance of Q_1. Now substituting for R_o from Eq. (13.96), we obtain

$$\text{CMRR} = g_{m1}\left(R_{o9} \parallel R_{o10}\right)/\epsilon_m \tag{13.102}$$

Before leaving this example, we observe that if the feedback were not present, the $2i$ term in Eq. (13.97) would be absent and the current i would become $\beta_P\left(v_{icm}/2R_o\right)$, which is β_P times higher than that when feedback is present. In other words, common-mode feedback reduces i, hence the common-mode transconductance and the common-mode gain, by a factor β_P. It can be shown that β_P is the magnitude of the loop gain. (See Exercise 13.17.)

EXERCISES

13.27 Show that if the source of the imbalance in the current-mirror load is that while $R_1 = R$, $R_2 = R + \Delta R$, the error ϵ_m is given by

$$\epsilon_m = \frac{\Delta R}{R + r_{e5} + \Delta R}$$

Evaluate ϵ_m for $\Delta R/R = 0.02$.
Ans. $\epsilon_m = 5.5 \times 10^{-3}$

13.28 Refer to Fig. 13.23 and assume that the bases of Q_9 and Q_{10} are at approximately constant voltages (signal ground). Find R_{o9}, R_{o10}, and hence R_o. Use $V_A = 125$ V for *npn* and 50 V for *pnp* transistors. Use the bias current values in Table 13.1.
Ans. $R_{o9} = 2.63\,\text{M}\Omega$; $R_{o10} = 31.1\,\text{M}\Omega$; $R_o = 2.43\,\text{M}\Omega$

13.29 Use the results of Exercises 13.27 and 13.28 to determine G_{mcm} and CMRR of the 741 input stage. What would the CMRR be if the common-mode feedback were not present? Assume $\beta_P = 50$.
Ans. $G_{mcm} = 1.13 \times 10^{-6}$ mA/V; CMRR $= 1.68 \times 10^5$ or 104.5 dB; without common-mode feedback, CMRR $= 70.5$ dB

The Second Stage Figure 13.24(a) shows the 741 second stage prepared for small-signal analysis, and Fig. 13.24(b) shows its small-signal model. The three model parameters R_{i2}, G_{m2}, and R_{o2} can be determined as follows.

The input resistance R_{i2} can be found by inspection to be

$$R_{i2} = (\beta_{16} + 1)\{r_{e16} + [R_9 \,\|\, (\beta_{17} + 1)(r_{e17} + R_8)]\} \tag{13.103}$$

From the equivalent circuit of Fig. 13.24(b), we see that the transconductance G_{m2} is the ratio of the *short-circuit output current* to the input voltage. Short-circuiting the output terminal of the second stage (Fig. 13.24a) to ground makes the signal current through the output resistance of Q_{13B} zero, and the output short-circuit current becomes equal to the collector signal current of Q_{17} (i_{c17}). This latter current can be easily related to v_{i2} as follows:

$$i_{c17} = \frac{\alpha v_{b17}}{r_{e17} + R_8} \tag{13.104}$$

$$v_{b17} = v_{i2} \frac{(R_9 \,\|\, R_{i17})}{(R_9 \,\|\, R_{i17}) + r_{e16}} \tag{13.105}$$

$$R_{i17} = (\beta_{17} + 1)(r_{e17} + R_8) \tag{13.106}$$

where we have neglected r_{o16} because $r_{o16} \gg R_9$. These equations can be combined to obtain

$$G_{m2} \equiv \frac{i_{c17}}{v_{i2}} \tag{13.107}$$

To determine the output resistance R_{o2} of the second stage in Fig. 13.24(a), we ground the input terminal and find the resistance looking back into the output terminal. It follows that R_{o2}

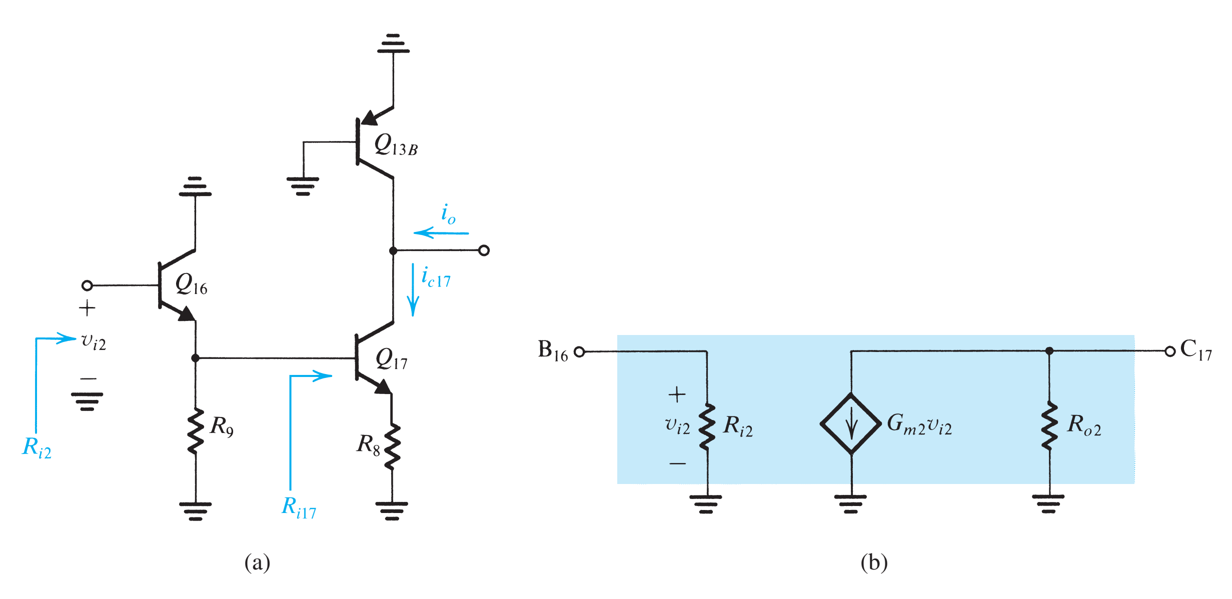


Figure 13.24 (**a**) The 741 second stage prepared for small-signal analysis. (**b**) Equivalent circuit.

is given by

$$R_{o2} = (R_{o13B} \| R_{o17}) \tag{13.108}$$

where R_{o13B} is the resistance looking into the collector of Q_{13B} while its base and emitter are connected to ground. It can be easily seen that

$$R_{o13B} = r_{o13B} \tag{13.109}$$

The second component in Eq. (13.108), R_{o17}, is the resistance seen looking into the collector of Q_{17}. Since the resistance between the base of Q_{17} and ground is relatively small (approximately equal to r_{e16}), one can considerably simplify matters by assuming that the base is grounded. Doing this, we can use Eq. (13.91) to determine R_{o17}.

EXERCISES

In the following exercises use $I_{C13B} = 550\ \mu\text{A}$, $I_{C16} = 16.2\ \mu\text{A}$, $I_{C17} = 550\ \mu\text{A}$, $\beta_N = 200$, $\beta_P = 50$, $V_{An} = 125$ V, $V_{Ap} = 50$ V, $R_9 = 50\ \text{k}\Omega$, and $R_8 = 100\ \Omega$.

13.30 Determine the value of R_{i2}.
Ans. 4 MΩ

13.31 Determine the value of G_{m2}.
Ans. 6.5 mA/V

13.32 Determine the values of R_{o13B}, R_{o17}, and R_{o2}.
Ans. 90.9 kΩ; 787 kΩ; 81 kΩ

13.33 Determine the value of the open-circuit voltage gain of the second stage.
Ans. −526.5 V/V

The Output Stage The 741 output stage is shown in Fig. 13.25 without the short-circuit-protection circuitry. The stage is shown driven by the second-stage transistor Q_{17} and loaded with a 2-kΩ resistance. The circuit is of the AB class (Section 12.4), with the network composed of Q_{18}, Q_{19}, and R_{10} providing the bias of the output transistors Q_{14} and Q_{20}. The use of this network rather than two diode-connected transistors in series enables biasing the output transistors at a low current (0.15 mA) in spite of the fact that the output devices are three times as large as the standard devices. This result is obtained by arranging that the current in Q_{19} is very small and thus its V_{BE} is also small. We analyzed the dc bias in Section 13.3.2.

Another feature of the 741 output stage worth noting is that the stage is driven by an emitter follower Q_{23}. As will be shown, this emitter follower provides added buffering, which makes the op-amp gain almost independent of the parameters of the output transistors.

Let's first determine the allowable range of output voltage swing. The maximum positive output voltage is limited by the saturation of current-source transistor Q_{13A}. Thus,

$$v_{O\max} = V_{CC} - |V_{CE\text{sat}}| - V_{BE14} \tag{13.110}$$

which is about 1 V below V_{CC}. The minimum output voltage (i.e., maximum negative amplitude) is limited by the saturation of Q_{17}. Neglecting the voltage drop across R_8, we obtain

$$v_{O\min} = -V_{EE} + V_{CE\text{sat}} + V_{EB23} + V_{EB20} \tag{13.111}$$

which is about 1.5 V above $-V_{EE}$.

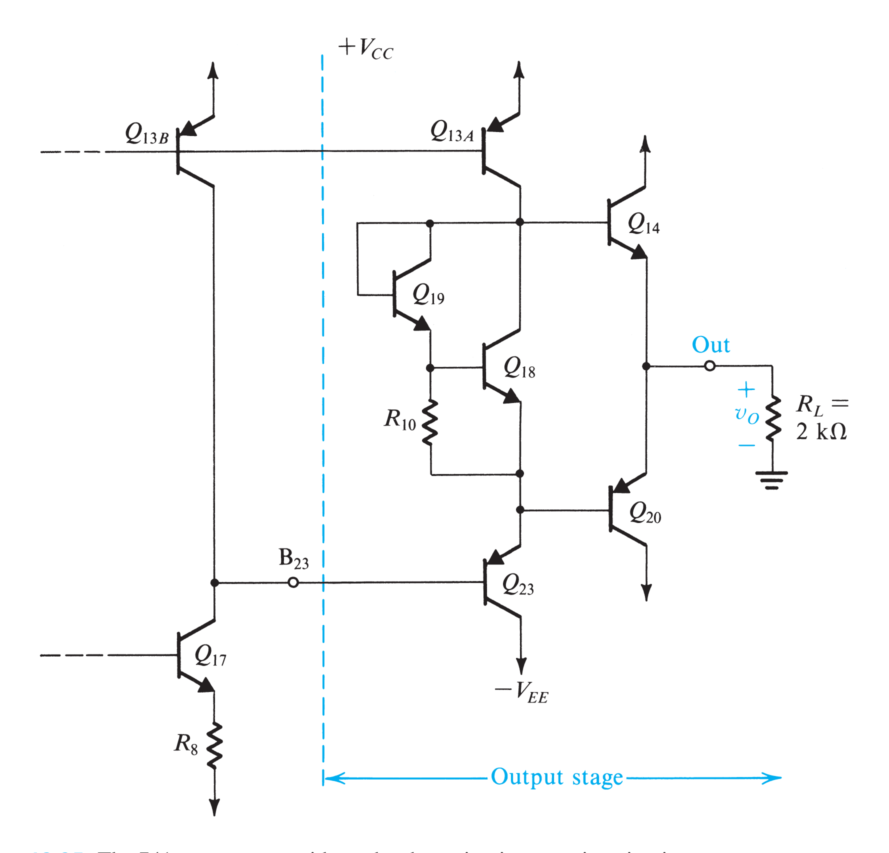


Figure 13.25 The 741 output stage without the short-circuit-protection circuitry.

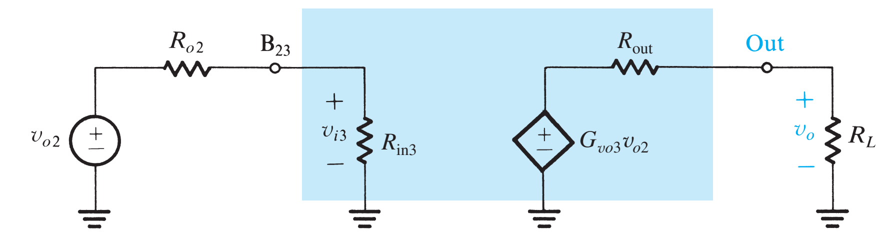


Figure 13.26 Model for the 741 output stage in Fig. 13.25.

Next, we consider the small-signal analysis of the output stage. Specifically, we model the output stage using the equivalent circuit in Fig. 13.26 and determine the model parameters as follows. Note that the model is shown fed with the open-circuit voltage of the second stage v_{o2}, where from Fig. 13.24(b), $v_{o2} = -G_{m2}R_{o2}v_{i2}$.

To determine the input resistance R_{in3}, we take into account the load resistance R_L and assume that one of the output transistors is conducting, as illustrated in the following example.

Example 13.6

Assuming that Q_{14} is off and Q_{20} is conducting a current of 5 mA to a load $R_L = 2\text{ k}\Omega$, determine the value of R_{in3}. Using $G_{m2} = 6.5$ mA/V and $R_{o2} = 81\text{ k}\Omega$, determine the voltage gain of the second stage.

Solution

Refer to Fig. 13.25. The input resistance looking into the base of Q_{20} is approximately $\beta_{20}R_L = 50 \times 2 = 100\text{ k}\Omega$. This resistance appears in parallel with the series combination of $r_{o13A} = V_{Ap}/I_{C13A} = 50\text{ V}/180\ \mu\text{A} = 280\text{ k}\Omega$, and the resistance of the Q_{18}–Q_{19} network. The latter resistance is very small (about 160 Ω; see later: Exercise 13.35). Thus, the total resistance in the emitter of Q_{23} is approximately $(100\text{ k}\Omega \,\|\, 280\text{ k}\Omega)$ or 74 kΩ, and the input resistance R_{in3} is obtained as

$$R_{in3} = \beta_{23} \times 74\text{ k}\Omega = 50 \times 74 = 3.7\text{ M}\Omega$$

We thus see that $R_{in3} \gg R_{o2}$, and the value of R_{in3} will have little effect on the performance of the op amp. Still we can determine the gain of the second stage as

$$A_2 \equiv \frac{v_{i3}}{v_{i2}} = -G_{m2}R_{o2}\frac{R_{in3}}{R_{in3}+R_{o2}}$$
$$= -6.5 \times 81\,\frac{3700}{3700+81} = -515\text{ V/V}$$

Continuing with the determination of the equivalent-circuit model parameters, we note from Fig. 13.26 that G_{vo3} is the **open-circuit overall voltage gain** of the output stage,

$$G_{vo3} = \left.\frac{v_o}{v_{o2}}\right|_{R_L=\infty} \tag{13.112}$$

With $R_L = \infty$, the gain of the emitter-follower output transistor (Q_{14} or Q_{20}) will be nearly unity. Also, with $R_L = \infty$ the resistance in the emitter of Q_{23} will be very large. This means that the gain of Q_{23} will be nearly unity and the input resistance of Q_{23} will be very large. We thus conclude that $G_{vo3} \simeq 1$.

Next, we shall find the value of the output resistance of the op amp, R_{out}. For this purpose refer to the circuit shown in Fig. 13.27. In accordance with the definition of R_{out} from Fig. 13.26, the input source feeding the output stage is grounded, but its resistance (which is the output resistance of the second stage, R_{o2}) is included. We have assumed that the output voltage v_O is negative, and thus Q_{20} is conducting most of the current; transistor Q_{14} has therefore been eliminated. The exact value of the output resistance will of course depend on which transistor (Q_{14} or Q_{20}) is conducting and on the value of load current. Nevertheless, we wish to find an estimate of R_{out}. The analysis for doing so is shown in Fig. 13.27. It should be noted, however, that to the value of R_{out} given in the figure we must add the resistance R_7 (27 Ω) (see Fig. 13.14), which is included for short-circuit protection, in order to obtain the total output resistance of the 741.

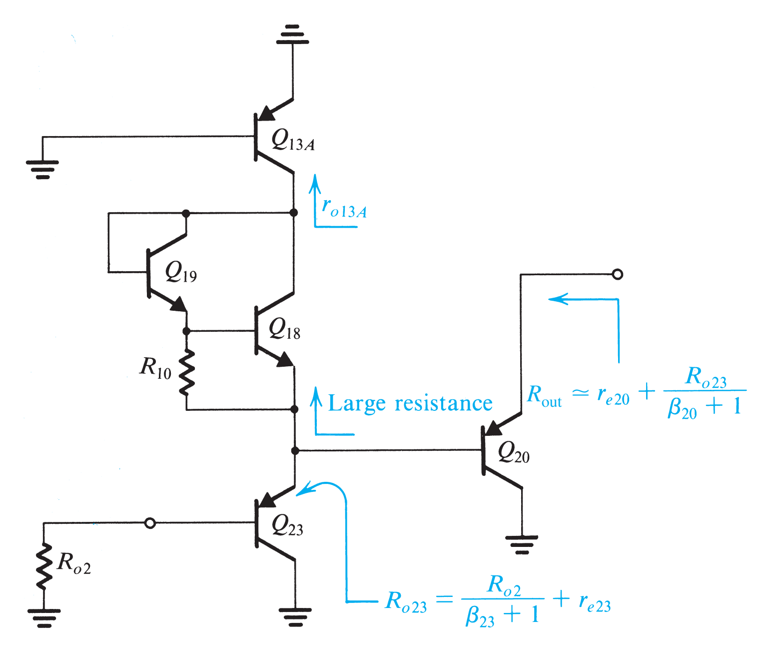


Figure 13.27 Circuit for finding the output resistance R_{out}.

EXERCISES

13.34 Find the value of R_{o23}, R_{out}, and the total output resistance of the 741 op amp. Use $R_{o2} = 81$ kΩ, $\beta_{23} = \beta_{20} = 50$, and $I_{C23} = 180$ µA, and assume that Q_{20} is conducting a load current of 5 mA.
Ans. 1.73 kΩ; 39 Ω; 66 Ω

13.35 Using a simple (r_π, g_m) model for each of the two transistors Q_{18} and Q_{19} in Fig. E13.35, find the small-signal resistance between A and A'. (*Note:* From Table 13.1, $I_{C18} = 165$ µA and $I_{C19} \simeq 16$ µA. Also, $\beta_N = 200$.)
Ans. 163 Ω

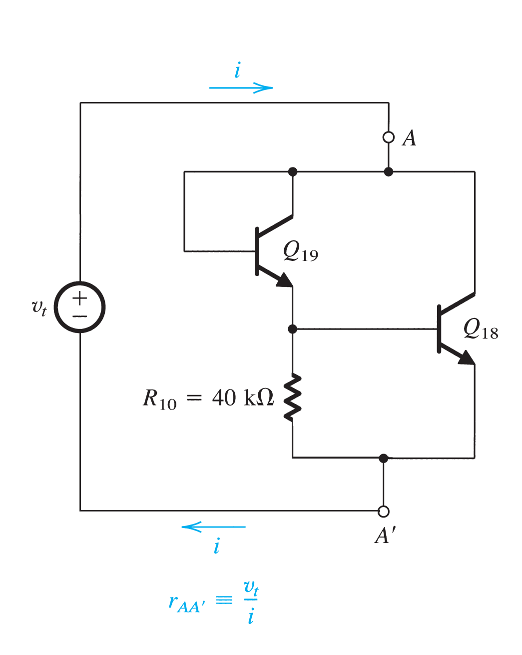


Figure E13.35

Output Short-Circuit Protection If the op-amp output terminal is short-circuited to one of the power supplies, one of the two output transistors could conduct a large amount of current. Such a large current can result in sufficient heating to cause burnout of the IC (Chapter 12). To guard against this possibility, the 741 op amp is equipped with a special circuit for short-circuit protection. The function of this circuit is to limit the current in the output transistors in the event of a short circuit.

Refer to Fig. 13.14 and note that the short-circuit-protection circuitry is highlighted in color. Resistance R_6 together with transistor Q_{15} limits the current that would flow out of Q_{14} in the event of a short circuit. Specifically, if the current in the emitter of Q_{14} exceeds about 20 mA, the voltage drop across R_6 exceeds 540 mV, which turns Q_{15} on. As Q_{15} turns on, its collector robs some of the current supplied by Q_{13A}, thus reducing the base current of Q_{14}. This mechanism thus limits the maximum current that the op amp can source (i.e., supply from the output terminal in the outward direction) to about 20 mA.

Limiting of the maximum current that the op amp can sink, and hence the current through Q_{20}, is done by a mechanism similar to the one discussed above. The relevant circuit is composed of R_7, Q_{21}, Q_{24}, and Q_{22}. For the components shown, the current in the inward direction is limited also to about 20 mA.

Overall Voltage Gain The overall small-signal gain can be found from the cascade of the equivalent circuits derived above for the three op-amp stages. This cascade is shown in Fig. 13.28, loaded with $R_L = 2\ \text{k}\Omega$, which is the typical value used in measuring

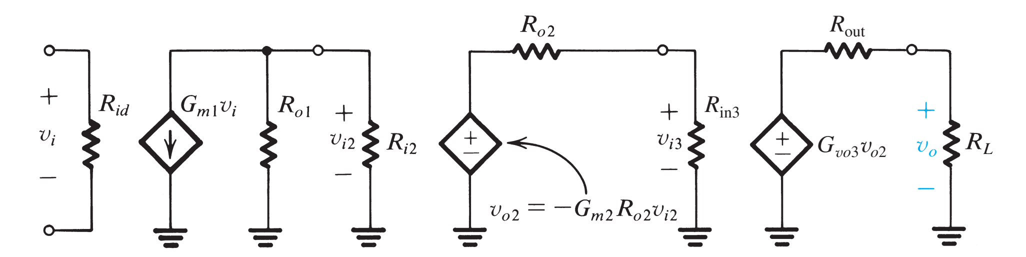


Figure 13.28 Cascading the small-signal equivalent circuits of the individual stages for the evaluation of the overall voltage gain.

and specifying the 741 data. The overall gain can be expressed as

$$\frac{v_o}{v_i} = \frac{v_{i2}}{v_i}\frac{v_{o2}}{v_{i2}}\frac{v_o}{v_{o2}} \tag{13.113}$$

$$= -G_{m1}(R_{o1} \parallel R_{i2})(-G_{m2}R_{o2})G_{vo3}\frac{R_L}{R_L + R_{\text{out}}} \tag{13.114}$$

Using the values found earlier yields for the overall open-circuit voltage gain,

$$A_0 \equiv \frac{v_o}{v_i} = -476.1 \times (-526.5) \times 0.97 = 243{,}147\text{ V/V} \tag{13.115}$$

$$= 107.7\text{ dB}$$

13.3.4 Frequency Response

The 741 is an internally compensated op amp. It employs the Miller compensation technique, studied in Section 11.10.3, to introduce a dominant low-frequency pole. Specifically, a 30-pF capacitor (C_C) is connected in the negative-feedback path of the second stage. An approximate estimate of the frequency of the dominant pole can be obtained as follows.

From Miller's theorem (Section 10.3.3), we see that the effective capacitance due to C_C between the base of Q_{16} and ground is (see Fig. 13.14)

$$C_{\text{in}} = C_C(1 + |A_2|) \tag{13.116}$$

where A_2 is the second-stage gain. Use of the value calculated for A_2 found in Example 13.6, $A_2 = -515$, results in $C_{\text{in}} = 15{,}480$ pF. Since this capacitance is quite large, we shall neglect all other capacitances between the base of Q_{16} and signal ground. The total resistance between this node and ground is

$$R_t = R_{o1} \parallel R_{i2}$$

$$= 6.7\,\text{M}\Omega \parallel 4\,\text{M}\Omega = 2.5\,\text{M}\Omega \tag{13.117}$$

Thus the dominant pole has a frequency f_P given by

$$f_P = \frac{1}{2\pi C_{\text{in}}R_t} = 4.1\text{ Hz} \tag{13.118}$$

It should be noted that this approach is equivalent to using the approximate formula in Eq. (11.93).

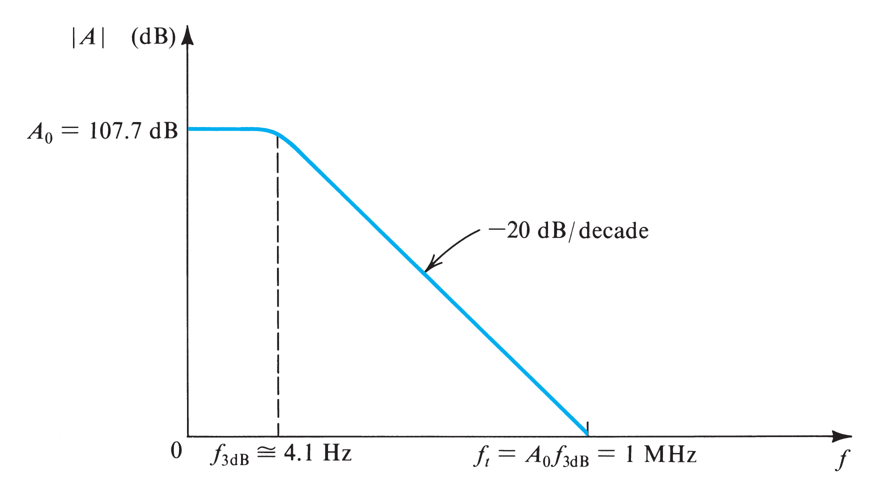


Figure 13.29 Bode plot for the 741 gain, neglecting nondominant poles.

As discussed in Section 11.10.3, Miller compensation provides an additional advantageous effect, namely, pole splitting. As a result, the other poles of the circuit are moved to very high frequencies. This has been confirmed by computer-aided analysis (see Gray et al., 2000).

Assuming that all nondominant poles are at very high frequencies, the calculated values give rise to the Bode plot shown in Fig. 13.29, where $f_{3dB} = f_P$. The unity-gain bandwidth f_t can be calculated from

$$f_t = A_0 f_{3dB} \tag{13.119}$$

Thus,

$$f_t = 243{,}147 \times 4.1 \simeq 1\text{ MHz} \tag{13.120}$$

Although this Bode plot implies that the phase shift at f_t is –90° and thus that the phase margin is 90°, in practice a phase margin of about 80° is obtained. The excess phase shift (about 10°) is due to the nondominant poles. This phase margin is sufficient to provide stable operation of closed-loop amplifiers with any value of feedback factor β. This convenience of use of the internally compensated 741 is achieved at the expense of a great reduction in open-loop gain and hence in the amount of negative feedback. In other words, if one requires a closed-loop amplifier with a gain of 1000, then the 741 is *overcompensated* for such an application, and one would be much better off designing one's own compensation (assuming, of course, the availability of an op amp that is not already internally compensated).

A Simplified Model The simplified model of the 741 op amp shown in Fig. 13.30 is similar to what we used for the CMOS two-stage op amp (Section 13.1.5). Here, however, the high-gain second stage, with its feedback capacitance C_C, is modeled by an ideal integrator. In this model, the gain of the second stage is assumed to be sufficiently large that a virtual ground appears at its input. For this reason the output resistance of the input stage and the input resistance of the second stage have been omitted. Furthermore, the output stage is assumed to be an ideal unity-gain follower. (Of course, the two-stage CMOS amplifier does not have an output stage.)

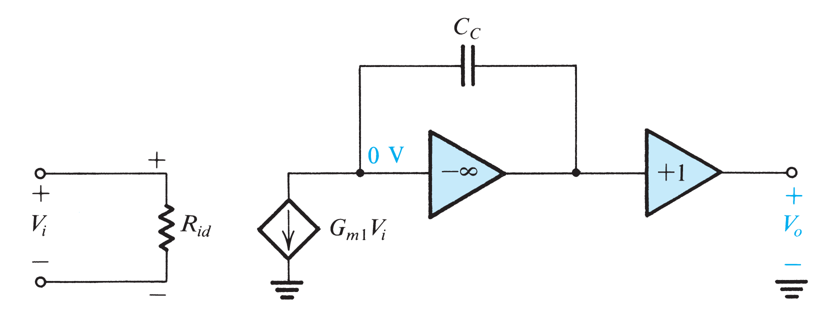


Figure 13.30 A simple model for the 741 based on modeling the second stage as an integrator.

Analysis of the model in Fig. 13.30 gives

$$A(s) \equiv \frac{V_o(s)}{V_i(s)} = \frac{G_{m1}}{sC_C} \tag{13.121}$$

Thus,

$$A(j\omega) = \frac{G_{m1}}{j\omega C_C} \tag{13.122}$$

and the magnitude of gain becomes unity at $\omega = \omega_t$, where

$$\omega_t = \frac{G_{m1}}{C_C} \tag{13.123}$$

Substituting $G_{m1} = 0.19$ mA/V and $C_C = 30$ pF yields

$$f_t = \frac{\omega_t}{2\pi} \simeq 1 \text{ MHz} \tag{13.124}$$

which is equal to the value calculated before. It should be pointed out, however, that this model is valid only at frequencies $f \gg f_{3dB}$. At such frequencies, the gain falls off with a slope of –20 dB/decade (Fig. 13.29), just like that of an integrator.

13.3.5 Slew Rate

The slew-rate limitation of op amps is discussed in Chapter 2, and expressions for *SR* are derived for the two-stage CMOS op amp in Section 13.1 and for the folded-cascode CMOS op amp in Section 13.2. The 741 slewing is very similar to that of the two-stage CMOS op amp. Thus, following an identical procedure, we can show that for the 741 op amp,

$$SR = \frac{2I}{C_C} \tag{13.125}$$

when $2I$ is the total bias current of the input differential stage.[7] For the 741, $I = 9.5$ μA, and $C_C = 30$ pF, resulting in $SR = 0.63$ V/μs.

[7]Note that in the CMOS two-stage op amp, the total bias current of the input stage is denoted I; hence the apparent discrepancy between the expressions in Eqs. (13.125) and (13.45).

Also, as we have done for the two-stage CMOS op amp, we can derive a relationship between SR and ω_t. For the 741 case, we can show that

$$SR = 4V_T\,\omega_t \tag{13.126}$$

where V_T is the thermal voltage (approximately 25 mV at room temperature). As a check, for the 741 we have

$$SR = 4 \times 25 \times 10^{-3} \times 2\pi \times 10^6 = 0.63\ \text{V}/\mu\text{s}$$

which is the result obtained previously. Observe that Eq. (13.126) is of the same form as Eq. (13.47), which applies to the two-stage CMOS op amp. Here, $4V_T$ replaces V_{OV}. Since, typically, V_{OV} will be two to three times the value of $4V_T$, a two-stage CMOS op amp with an f_t equal to that of the 741 exhibits a slew rate that is two to three times as large as that of the 741.

EXERCISE

13.36 Use the value of the slew rate calculated above to find the full-power bandwidth f_M of the 741 op amp. Assume that the maximum output is ±10 V.
Ans. 10 kHz



13.4 Modern Techniques for the Design of BJT Op Amps

Although the ingenious techniques employed in the design of the 741 op amp have stood the test of time, they are now more than 40 years old! Technological advances have resulted in changes in the user requirements of general-purpose bipolar op amps. The resulting more demanding specifications have in turn posed new challenges to analog IC designers, who, as they have done repeatedly before, are responding with new and exciting circuits. In this section we present a sample of recently developed design techniques. For more on this rather advanced topic the reader is referred to the Analog Circuits section of the bibliography in Appendix I.

13.4.1 Special Performance Requirements

Many of the special performance requirements stem from the need to operate modern op amps from power supplies of much lower voltages. Thus while the 741-type op amp operated from ±15-V power supplies, many modern BJT op amps are required to operate from a *single power supply of only 2* V *to 3* V. This is done for a number of reasons, including the following.

1. Modern small-feature-size IC fabrication technologies require low power-supply voltages.
2. Compatibility must be achieved with other parts of the system that use low-voltage supplies.
3. Power dissipation must be minimized, especially for battery-operated equipment.

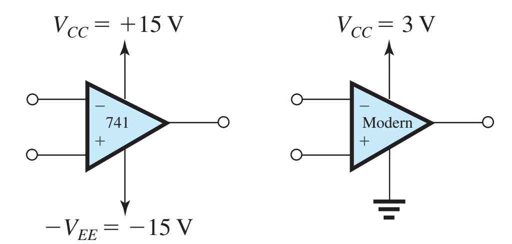


Figure 13.31 Power-supply requirements have changed considerably. Modern BJT op amps are required to operate from a single supply V_{CC} of 2 V to 3 V.

As Fig. 13.31 indicates, there are two important changes: the use of a single ground-referenced power supply V_{CC}, and the low value of V_{CC}. Both of these requirements give rise to changes in performance specifications and pose new design challenges. In the following we discuss two of the resulting changes.

Rail-to-Rail Input Common-Mode Range Recall that the input common-mode range of an op amp is the range of common-mode input voltages for which the op amp operates properly and meets its performance specifications, such as voltage gain and CMRR. Op amps of the 741 type operate from ±15-V supplies and exhibit an input common-mode range that extends to within a couple of volts of each supply. Such a gap between the input common-mode range and the power supply is obviously unacceptable if the op amp is to be operated from a single supply that is only 2 V to 3 V. Indeed we will now show that these single-supply, low-voltage op amps need to have an input common-mode range that extends over the entire supply voltage, 0 to V_{CC}, referred to as rail-to-rail input common-mode range.

Consider first the inverting op-amp configuration shown in Fig. 13.32(a). Since the positive input terminal is connected to ground (which is the voltage of the negative supply rail), ground voltage has to be within the allowable input common-mode range. In fact, because for positive output voltages the voltage at the inverting input terminal can go slightly negative, the input common-mode range should extend below the negative supply rail (ground).

Next consider the unity-gain voltage follower obtained by applying 100% negative feedback to an op amp, as shown in Fig. 13.32(b). Here the input common-mode voltage is equal to the input signal v_I. To maximize the usefulness of this buffer amplifier, its input signal v_I should be allowed to extend from 0 to V_{CC}, especially since V_{CC} is only 2 V to 3 V. Thus the input common-mode range should include also the positive supply rail. As will be seen shortly, modern BJT op amps can operate over an input common-mode voltage range that extends a fraction of a volt beyond its two supply rails: that is, more than rail-to-rail operation!

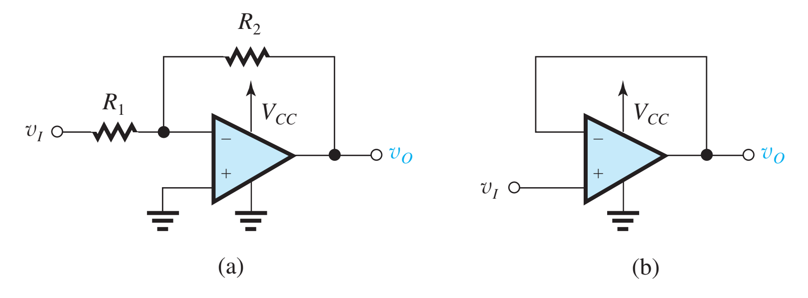


Figure 13.32 **(a)** In the inverting configuration, the positive op-amp input is connected to ground; thus it is imperative that the input common-mode range includes ground voltage. **(b)** In the unity-gain follower configuration, $v_{ICM} = v_I$; thus it is highly desirable for the input common-mode range to include ground voltage and V_{CC}.

Near Rail-to-Rail Output Signal Swing In the 741 op amp, we were satisfied with an output that can swing to within 2 V or so of each of the supply rails. With a supply of ±15 V, this capacity resulted in a respectable ±13-V output range. However, to limit the output swing to within 2 V of the supply rails in an op amp operating from a single 3-V supply would result in an unusable device! Thus, here too, we require near rail-to-rail operation. As we shall see in Section 13.4.5, this requirement forces us to adopt a whole new approach to output-stage design.

Device Parameters The technology we shall use in the examples, exercises, and problems for this section has the following characteristics:

$$\begin{array}{lll} npn \text{ transistors}: & \beta = 40 & V_A = 30\ V \\ pnp \text{ transistors}: & \beta = 10 & |V_A| = 20\ V \end{array}$$

For both, $|V_{BE}| \simeq 0.7$ V and $|V_{CEsat}| \simeq 0.1$ V. It is important to note that we will assume that for this technology, the transistor will remain in the active mode for $|V_{CE}|$ as low as 0.1 V (in other words, that 0.6 V is needed to forward-bias the CBJ).

13.4.2 Bias Design

As in the 741 circuit, the bias design of modern BJT amplifiers makes extensive use of current mirrors and current-steering circuits (Sections 8.2 and 8.6). Typically, however, the bias currents are small (in the microamp range). Thus, the Widlar current source (Section 8.6.4) is especially popular here. As well, emitter-degeneration resistors (in the tens-of-kilohm range) are frequently used.

Figure 13.33 shows a self-biased current-reference source that utilizes a Widlar circuit formed by Q_1, Q_2, and R_2, and a current mirror Q_3–Q_4 with matched emitter-degeneration resistors R_3 and R_4. The circuit establishes a current I in each of the four transistors, with the value of I determined as follows. Neglecting base currents and r_o's for simplicity, we write

$$V_{BE1} = V_T \ln\left(\frac{I}{I_{S1}}\right)$$

$$V_{BE2} = V_T \ln\left(\frac{I}{I_{S2}}\right)$$

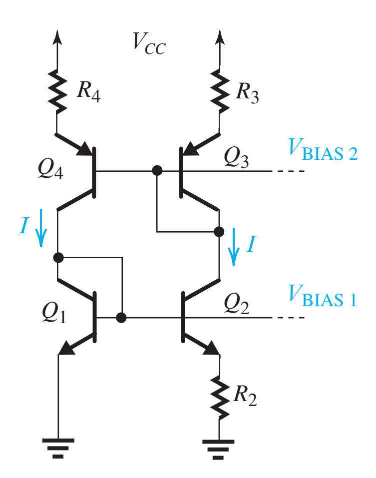


Figure 13.33 A self-biased current-reference source utilizing a Widlar circuit to generate $I = (V_T/R_2)\ln(I_{S2}/I_{S1})$. The bias voltages V_{BIAS1} and V_{BIAS2} are utilized in other parts of the op-amp circuit for biasing other transistors.

Thus,

$$V_{BE1} - V_{BE2} = V_T \ln\left(\frac{I_{S2}}{I_{S1}}\right)$$

But,

$$V_{BE1} - V_{BE2} = IR_2$$

Thus,

$$I = \frac{V_T}{R_2} \ln\left(\frac{I_{S2}}{I_{S1}}\right) \tag{13.127}$$

Thus the value of I is determined by R_2 and the ratio of the emitter areas of Q_1 and Q_2. Also, observe that I is independent of V_{CC}, a highly desirable outcome. Neglecting the temperature dependence of R_2, we see that I is directly PTAT (*p*roportional *t*o the *a*bsolute *t*emperature *T*). It follows that transistors biased by I or mirrored versions of it will exhibit g_m's that are constant independent of temperature!

EXERCISE

D13.37 Design the circuit in Fig. 13.33 to generate a current $I = 10\ \mu A$. Utilize transistors Q_1 and Q_2 having their areas in a 1:2 ratio. Assume that Q_3 and Q_4 are matched and design for a 0.2-V drop across each of R_3 and R_4. Specify the values of R_2, R_3, and R_4.
Ans. 1.73 kΩ; 20 kΩ; 20 kΩ

The circuit in Fig. 13.33 provides a bias line V_{BIAS1} with a voltage equal to V_{BE1}. This can be used to bias other transistors and thus generate currents proportional to I by appropriately scaling their emitter areas. Similarly, the circuit provides a bias line V_{BIAS2} at a voltage $(IR_3 + V_{EB3})$ below V_{CC}. This bias line can be used to bias other transistors and thus generate constant currents proportional to I by appropriately scaling emitter areas and emitter-degeneration resistances. These ideas are illustrated in Fig. 13.34.

EXERCISE

D13.38 Refer to the circuit in Fig. 13.34 and assume that the V_{BIAS2} line is connected to the corresponding line in Fig. 13.33. It is required to generate currents $I_8 = 10\ \mu A$, $I_9 = 20\ \mu A$, and $I_{10} = 5\ \mu A$. Specify the required emitter areas of Q_8, Q_9, and Q_{10} as ratios of the emitter area of Q_3. Also specify the values required for R_8, R_9, and R_{10}. Use the values of R_3 and R_4 found in Exercise 13.37. Ignore base currents.
Ans. 1, 2, 0.5; 20 kΩ, 10 kΩ, 40 kΩ

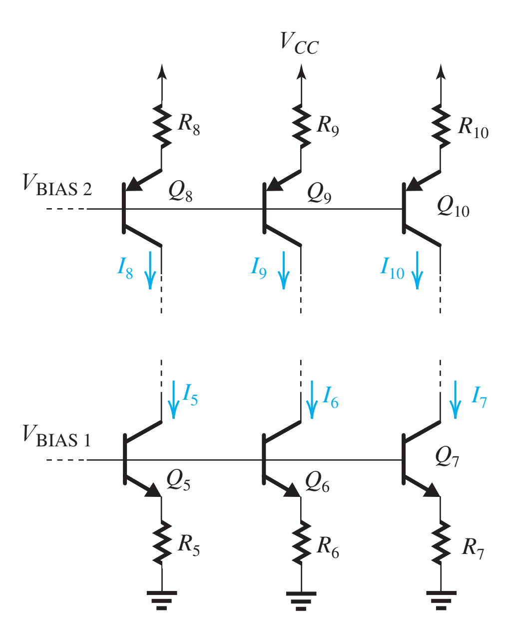


Figure 13.34 The bias lines V_{BIAS1} and V_{BIAS2} provided by the circuit in Fig. 13.33 are utilized to bias other transistors and generate constant currents I_5 to I_{10}. Both the transistor area and the emitter-degeneration resistance value have to be appropriately scaled.

13.4.3 Design of the Input Stage to Obtain Rail-to-Rail V_{ICM}

The classical differential input stage with current-mirror load is shown in Fig. 13.35(a). This is essentially the core of the 741 input stage, except that here we are using a single positive power supply. As well, the CMOS counterpart of this circuit is utilized in nearly every CMOS op-amp design (see Section 13.1). Unfortunately, this very popular circuit does not meet our requirement of rail-to-rail common-mode operation.

Consider first the low end of the input common-mode range. The value of $V_{ICM\min}$ is limited by the need to keep Q_1 in the active mode. Specifically, since the collector of Q_1 is at a voltage $V_{BE3} \simeq 0.7$ V, we see that the voltage applied to the base of Q_1 cannot go lower than 0.1 V without causing the collector–base junction of Q_1 to become forward biased.

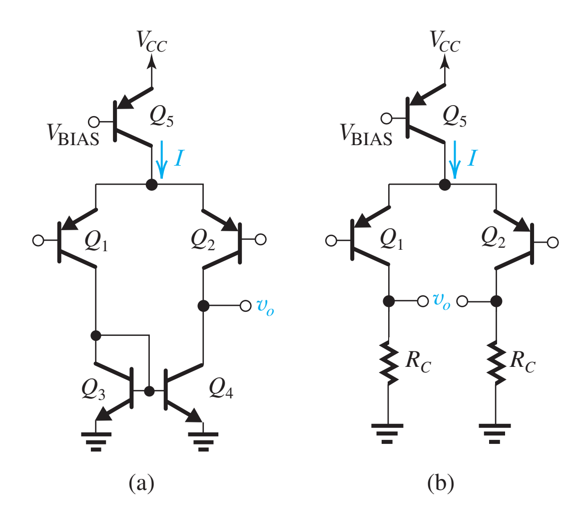


Figure 13.35 For the input common-mode range to include ground voltage, the classical current-mirror-loaded input stage in (**a**) has to be replaced with the resistively loaded configuration in (**b**) with the dc voltage drop across R_C limited to 0.2 V to 0.3 V.

Thus $V_{ICM\min} = 0.1$ V, and the input common-mode range does *not* include ground voltage as required.

The only way to extend $V_{ICM\min}$ to 0 V is to lower the voltage at the collector of Q_1. This in turn can be achieved only by abandoning the use of the current-mirror load and utilizing instead resistive loads, as shown in Fig. 13.35(b). Observe that in effect we are *going back* to the resistively loaded differential pair with which we began our study of differential amplifiers in Chapter 9!

The minimum allowed value of V_{ICM} in the circuit of Fig. 13.35(b) is still of course limited by the need to keep Q_1 and Q_2 in the active mode. This in turn is achieved by avoiding V_{ICM} values that cause the base voltages of Q_1 and Q_2 to go below their collector voltages by more than 0.6 V,

$$V_{ICM\min} = V_{R_C} - 0.6 \text{ V}$$

where V_{R_C} is the voltage drop across each of R_{C1} and R_{C2}. Now if V_{R_C} is selected to be 0.2 V to 0.3 V, then $V_{ICM\min}$ will be −0.4 V to −0.3 V, which is exactly what we need.

The major drawback of replacing the current-mirror load with resistive loads is that the differential gain realized is considerably reduced,

$$\frac{v_o}{v_{id}} = -g_{m1,2}R_C$$

$$= -\frac{I/2}{V_T}R_C = -\frac{V_{R_C}}{V_T}$$

where we have neglected r_o for simplicity. Thus for $V_{R_C} = 0.3$ V, the gain realized is only 12 V/V. As we will see shortly, this low-gain problem can be solved by cascoding.

Next consider the upper end of the input common-mode range. Reference to the circuit in Fig. 13.35(b) shows that the maximum voltage that can be applied to the bases of Q_1 and Q_2 is limited by the need to keep the current-source transistor in the active mode. This in turn is achieved by ensuring that the voltage across Q_5, V_{EC5}, does not fall below 0.1 V or so. Thus the maximum value of V_{ICM} will be a voltage $V_{EB1,2}$ or approximately 0.7 V lower,

$$V_{ICM\max} = V_{CC} - 0.1 - 0.7 = V_{CC} - 0.8$$

That is, the upper end of the input common-mode range is at least 0.8 V below V_{CC}, a severe limitation.

To recap, while the circuit in Fig. 13.35(b) has $V_{ICM\min}$ of a few tenths of a volt below the negative power-supply rail (at ground voltage), the upper end of V_{ICM} is rather far from V_{CC},

$$-0.3 \leq V_{ICM} \leq V_{CC} - 0.8$$

where we have assumed $V_{R_C} = 0.3$ V. To extend the upper end of V_{ICM}, we adopt a solution similar to that used in the CMOS case (Section 13.2.6, Fig. 13.12), namely, we utilize a parallel complementary input stage. Toward that end, note that the *npn* version of the circuit of Fig. 13.35(b), shown in Fig. 13.36, has a common-input range of

$$0.8 \leq V_{ICM} \leq V_{CC} + 0.3$$

where we have assumed that $V_{R_C} = 0.3$ V. Thus, as expected, the high end meets our specifications and in fact is above the positive supply rail by 0.3 V. The lower end, however, does not, but this should cause us no concern because the lower end will be looked after by the

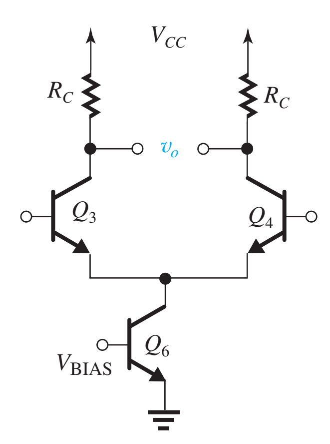


Figure 13.36 The complement of the circuit in Fig. 13.35(b). While the input common-mode range of the circuit in Figure 13.35(b) extends below ground, here it extends above V_{CC}. Connecting the two circuits in parallel, as will be shown, results in a rail-to-rail V_{ICM} range.

pnp pair. Finally, note that there is a range of V_{ICM} in which both the *pnp* and the *npn* circuits will be active and properly operating,

$$0.8 \leq V_{ICM} \leq V_{CC} - 0.8$$

Figure 13.37 shows an input stage that achieves more than rail-to-rail input common-mode range by utilizing a *pnp* differential pair (Q_1, Q_2) and an *npn* differential pair (Q_3, Q_4), connected in parallel. To keep the diagram simple, we are not showing the parallel connection of the input terminals; the + input terminals are assumed to be connected together, and similarly for the – input terminals. In order to increase the gain obtained from the resistively loaded differential pairs, a folded-cascode stage is added. Here R_7 and R_8 are the resistive loads of the *pnp* pair Q_1-Q_2, and Q_7-Q_8 are its cascode transistors. Similarly, R_9 and R_{10} are

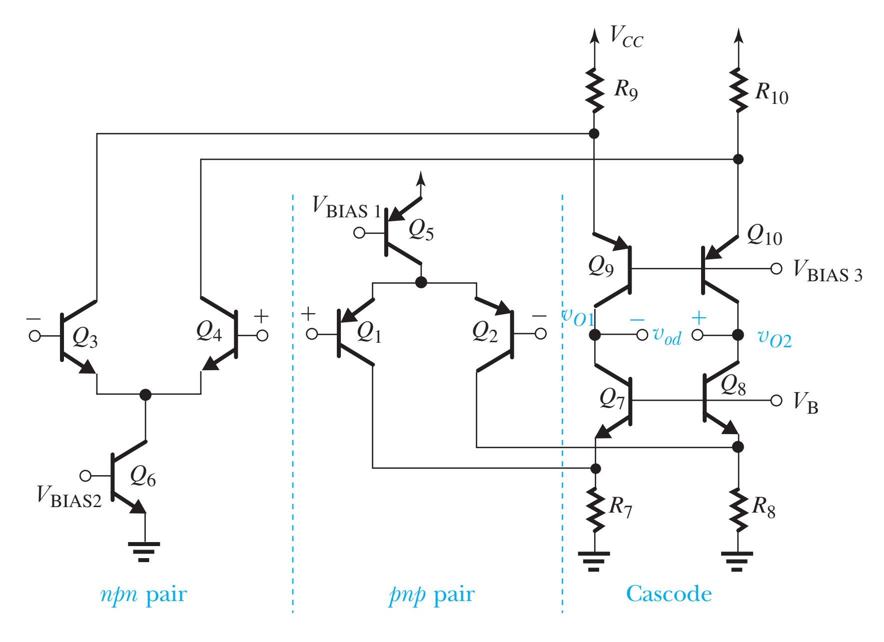


Figure 13.37 Input stage with rail-to-rail input common-mode range and a folded-cascode stage to increase the gain. Note that all the bias voltages including V_{BIAS3} and V_B are generated elsewhere on the chip.

the resistive loads of the *npn* pair Q_3-Q_4, and Q_9-Q_{10} are its cascode transistors. Observe that the cascode transistors do "double duty." For instance, Q_7-Q_8 operate as the cascode devices for Q_1-Q_2 and at the same time as current-source loads for Q_9-Q_{10}. A similar statement can be made about Q_9-Q_{10}. The output voltage of the first stage, v_{od}, is taken between the collectors of the cascode devices.

For $V_{ICM} \ll 0.8$ V, the *npn* stage will be inactive and the gain is determined by the transconductance G_m of the Q_1-Q_2 pair together with the output resistance seen between the collectors of the cascode transistors. At the other end of V_{ICM}, that is, $V_{ICM} \gg V_{CC} - 0.8$, the Q_1-Q_2 stage will be inactive, and the gain will be determined by the transconductance G_m of the Q_3-Q_4 pair and the output resistance between the collectors of the cascode devices. In the overlap region $0.8 \leq V_{ICM} \leq V_{CC} - 0.8$, both the *pnp* and *npn* stages will be active and their effective transconductances G_m add up, thus resulting in a higher gain. The dependence of the differential gain on the input common-mode V_{ICM} is usually undesirable and can be reduced considerably by arranging that one of the two differential pairs is turned off when the other one is active.[8]

Example 13.7

It is required to find the input resistance and the voltage gain of the input stage shown in Fig. 13.37. Let $V_{ICM} \ll 0.8$ V so that the Q_3-Q_4 pair is off. Assume that Q_5 supplies 10 μA, that each of Q_7 to Q_{10} is biased at 10 μA, and that all four cascode transistors are operating in the active mode. The input resistance of the second stage of the op amp (not shown) is $R_L = 2\,\text{M}\Omega$. The emitter-degeneration resistances are $R_7 = R_8 = 20\,\text{k}\Omega$, and $R_9 = R_{10} = 30\,\text{k}\Omega$. Recall that the device parameters are $\beta_N = 40$, $\beta_P = 10$, $V_{An} = 30$ V, $|V_{Ap}| = 20$ V.

Solution

Since the stage is fully balanced, we can use the differential half-circuit shown in Fig. 13.38(a). The input resistance R_{id} is twice the value of $r_{\pi 1}$,

$$R_{id} = 2r_{\pi 1} = 2\beta_P/g_{m1}$$

where

$$g_{m1} = \frac{I_{C1}}{V_T} = \frac{5 \times 10^{-6}}{25 \times 10^{-3}} = 0.2\ \text{mA/V}$$

Thus,

$$R_{id} = \frac{2 \times 10}{0.2} = 100\ \text{k}\Omega$$

To find the short-circuit transconductance, we short the output to ground as shown in Fig. 13.38(b) and find G_{m1} as

$$G_{m1} = \frac{i_o}{v_{id}/2}$$

[8]This is done in the NE5234 op amp, whose circuit is described and analyzed in great detail in Gray et al. (2009).

Example 13.7 *continued*

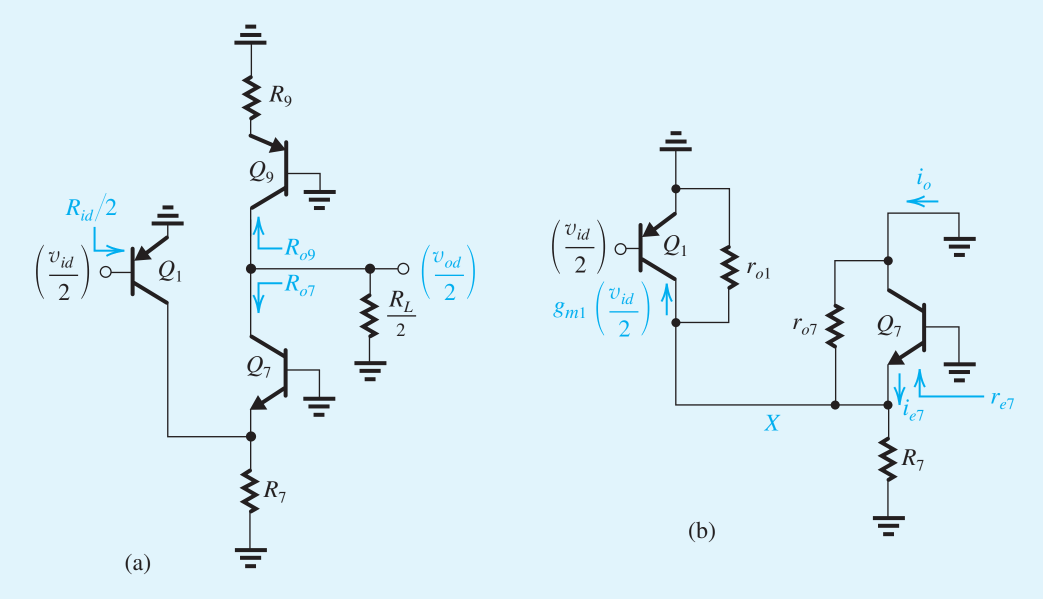


Figure 13.38 **(a)** Differential half-circuit for the input stage shown in Fig. 13.37 with $V_{ICM} \ll 0.8$ V. **(b)** Determining $G_{m1} = i_o/(v_{id}/2)$

At node X we have four parallel resistances to ground,

$$r_{o1} = \frac{|V_{Ap}|}{I_{C1}} = \frac{20 \text{ V}}{5 \ \mu\text{A}} = 4 \text{ M}\Omega$$

$$R_7 = 20 \text{ k}\Omega$$

$$r_{o7} = \frac{V_{An}}{I_{C7}} = \frac{30 \text{ V}}{10 \ \mu\text{A}} = 3 \text{ M}\Omega$$

$$r_{e7} \simeq \frac{1}{g_{m7}} = \frac{V_T}{I_{C7}} = \frac{25 \text{ mV}}{10 \ \mu\text{A}} = 2.5 \text{ k}\Omega$$

Obviously r_{o1} and r_{o7} are very large and can be neglected. Then, the portion of $g_{m1}(v_{id}/2)$ that flows into the emitter proper of Q_7 can be found from

$$i_{e7} \simeq g_{m1}\left(\frac{v_{id}}{2}\right)\frac{R_7}{R_7 + r_{e7}}$$

$$= g_{m1}\left(\frac{v_{id}}{2}\right)\frac{20}{20 + 2.5} = 0.89 g_{m1}\left(\frac{v_{id}}{2}\right)$$

and the output short-circuit current i_o is

$$i_o \simeq i_{e7} = 0.89 g_{m1}(v_{id}/2)$$

Thus,

$$G_{m1} \equiv \frac{i_o}{v_{id}/2} = 0.89 g_{m1} = 0.89 \times 0.2 = 0.18 \text{ mA/V}$$

To find the voltage gain we need to determine the total resistance between the output node and ground for the circuit in Fig. 13.38(a),

$$R = R_{o9} \parallel R_{o7} \parallel (R_L/2)$$

The resistance R_{o9} is the output resistance of Q_9, which has an emitter-degeneration resistance R_9. Thus R_{o9} can be found using Eq. (8.68),

$$R_{o9} = r_{o9} + (R_9 \parallel r_{\pi 9})(1 + g_{m9} r_{o9})$$

where

$$r_{o9} = \frac{|V_{Ap}|}{I_{C9}} = \frac{20 \text{ V}}{10 \text{ μA}} = 2 \text{ MΩ}$$

$$g_{m9} = \frac{I_{C9}}{V_T} = \frac{10 \text{ μA}}{25 \text{ mV}} = 0.4 \text{ mA/V}$$

$$r_{\pi 9} = \frac{\beta_P}{g_{m9}} = \frac{10}{0.4 \text{ mA/V}} = 25 \text{ kΩ}$$

Thus

$$R_{o9} = 2 + (30 \parallel 25) \times 10^{-3} \left(1 + 0.4 \times 2 \times 10^3\right)$$
$$= 12.9 \text{ MΩ}$$

The resistance R_{o7} is the output resistance of Q_7, which has an emitter-degeneration resistance $(R_7 \parallel r_{o1}) \simeq R_7$. Thus,

$$R_{o7} = r_{o7} + (R_7 \parallel r_{\pi 7})(1 + g_{m7} r_{o7})$$

where

$$r_{o7} = \frac{V_{An}}{I_{C7}} = \frac{30 \text{ V}}{10 \text{ μA}} = 3 \text{ MΩ}$$

$$g_{m7} = \frac{I_{C7}}{V_T} = \frac{10 \text{ μA}}{25 \text{ mV}} = 0.4 \text{ mA/V}$$

$$r_{\pi 7} = \frac{\beta_N}{g_{m7}} = \frac{40}{0.4} = 100 \text{ kΩ}$$

Thus,

$$R_{o7} = 3 + (20 \parallel 100) \times 10^{-3} \left(1 + 0.4 \times 3 \times 10^3\right)$$
$$= 23 \text{ MΩ}$$
$$\frac{R_L}{2} = \frac{2 \text{ MΩ}}{2} = 1 \text{ MΩ}$$

Example 13.7 *continued*

The total resistance R can now be found as

$$R = 12.9 \| 23 \| 1 = 0.89\,\text{M}\Omega$$

Finally, we can find the voltage gain as

$$A_d = \frac{v_{od}/2}{v_{id}/2} = G_{m1}R$$
$$= 0.18 \times 0.89 \times 10^3 = 160\,\text{V/V}$$

13.4.4 Common-Mode Feedback to Control the DC Voltage at the Output of the Input Stage

For the cascode circuit in Fig. 13.37 to operate properly and provide high output resistance and thus high voltage gain, the cascode transistors Q_7 through Q_{10} must operate in the active mode at all times. However, relying solely on matching will not be sufficient to ensure that the currents supplied by Q_9 and Q_{10} are exactly equal to the currents supplied by Q_7 and Q_8. Any small mismatch ΔI between the two sets of currents will be multiplied by the large output resistance between each of the collector nodes and ground, and thus there will be large changes in the voltages v_{O1} and v_{O2}. These changes in turn can cause one set of the current sources (i.e., Q_7–Q_8 or Q_9–Q_{10}) to saturate. We therefore need a circuit that detects the change in the dc or common-mode component V_{CM} of v_{O1} and v_{O2},

$$V_{CM} = \frac{1}{2}(v_{O1} + v_{O2}) \tag{13.128}$$

and adjusts the bias voltage on the bases of Q_7 and Q_8, V_B, to restore current equality. This negative-feedback loop should be insensitive to the differential signal components of v_{O1} and v_{O2}; otherwise it would reduce the differential gain. Thus the feedback loop should provide **common-mode feedback** (CMF).

Figure 13.39 shows the cascode circuit with the CMF circuit as a black box. The CMF circuit accepts v_{O1} and v_{O2} as inputs and provides the bias voltage V_B as output. In a particular implementation we will present shortly, the CMF circuit has the transfer characteristic

$$V_B = V_{CM} + 0.4 \tag{13.129}$$

By keeping V_B higher than V_{CM} by only 0.4 V, the CMF circuit ensures that Q_7 and Q_8 remain active (0.6 V is needed for saturation).

The nominal value of V_B is determined by the quiescent current of Q_7 through Q_{10}, the quiescent value of I_1 and I_2, and the value of R_7 and R_8. The resulting nominal value of V_B and the corresponding value of V_{CM} from Eq. (13.129) are designed to ensure that Q_9 and Q_{10} operate in the active mode. Here, it is important to recall that V_{BIAS3} is determined by the rest of the op-amp bias circuit.

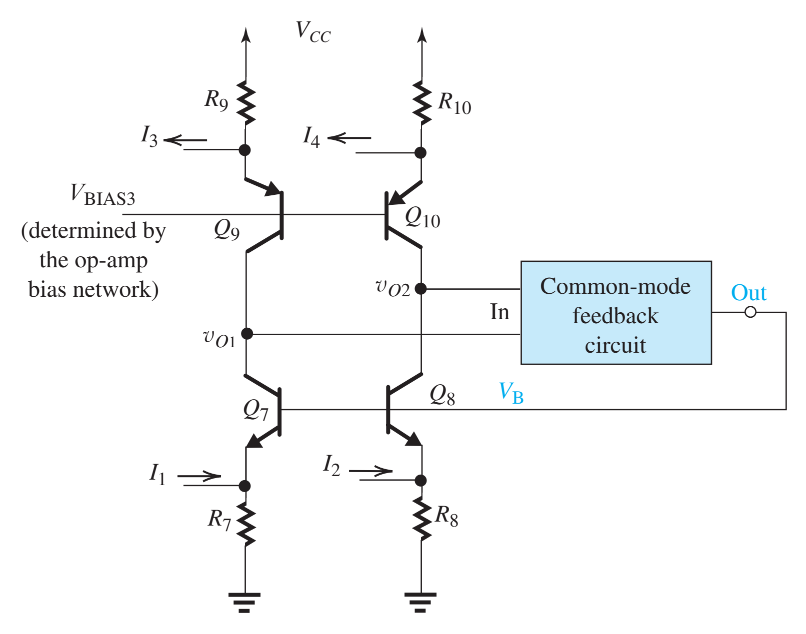


Figure 13.39 The cascode output circuit of the input stage and the CMF circuit that responds to the common-mode component $V_{CM} = \frac{1}{2}(v_{O1} + v_{O2})$ by adjusting V_B so that Q_7–Q_8 conduct equal currents to Q_9–Q_{10}, and Q_7–Q_{10} operate in the active mode.

To see how the CMF circuit regulates the dc voltage V_{CM}, assume that for some reason V_B is higher than it should be and as a result the currents of Q_7 and Q_8 exceed the currents supplied by Q_9 and Q_{10} by an increment ΔI. When multiplied by the total resistance between each of the output nodes and ground, the increment ΔI will result in a large negative voltage increment in v_{O1} and v_{O2}. The CMF circuit responds by lowering V_B to the value that restores the equality of currents. The change in V_B needed to restore equilibrium is usually small (see Example 13.8 below) and according to Eq. (13.129) the corresponding change in V_{CM} will be equally small. Thus we see negative feedback in action: It minimizes the initial change and thus keeps V_{CM} nearly constant at its nominal value, which is designed to operate Q_7 through Q_{10} in the active region.

We conclude by considering briefly a possible implementation of the CMF circuit. Figure 13.40 shows the second stage of an op-amp circuit. The circuit is fed by the outputs of the input stage, v_{O1} and v_{O2},

$$v_{O1} = V_{CM} + v_d/2$$
$$v_{O2} = V_{CM} - v_d/2$$

In addition to amplifying the differential component of v_d, the circuit generates a dc voltage V_B,

$$V_B = V_{CM} + 0.4$$

To see how the circuit works, note that Q_{11} and Q_{12} are emitter followers that minimize the loading of the second stage on the input stage. The emitter followers deliver to the bases of the differential pair Q_{13}–Q_{14} voltages that are almost equal to v_{O1} and v_{O2} but dc shifted

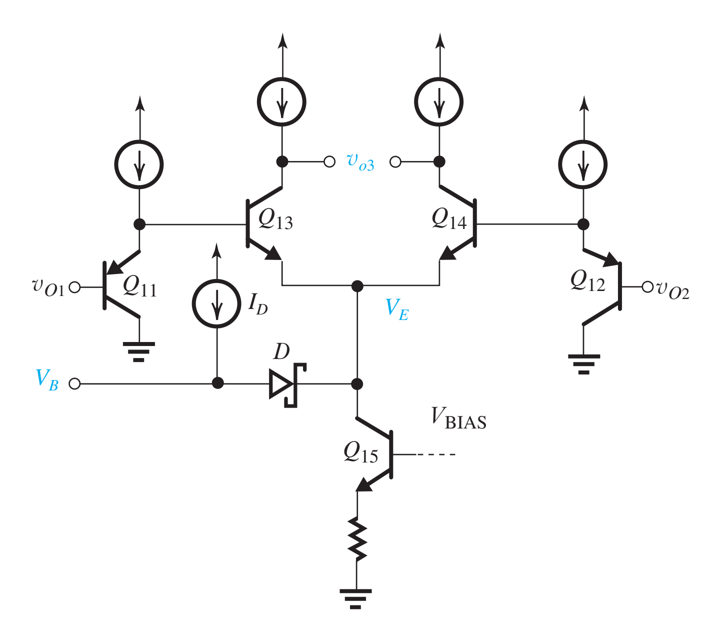


Figure 13.40 An op-amp second stage incorporating the common-mode feedback circuit for the input stage. Note that the circuit generates the voltage V_B needed to bias the cascode circuit in the first stage. Diode *D* is a Schottky-barrier diode, which exhibits a forward voltage drop of about 0.4 V.

by $V_{EB11,12}$. Thus the voltage at the emitters of Q_{13}–Q_{14} will be

$$V_E = V_{CM} + V_{EB11,12} - V_{BE13,14}$$

which reduces to

$$V_E \simeq V_{CM}$$

The voltage V_B is simply equal to V_E plus the voltage drop of diode D_1. The latter is a **Schottky-barrier diode** (SBD), which features a low forward drop of about 0.4 V. Thus,

$$V_B = V_E + V_D = V_{CM} + 0.4$$

as required.

Example 13.8

Consider the operation of the circuit in Fig. 13.39. Assume that $V_{ICM} \ll 0.8$ V and thus the *npn* input pair (Fig. 13.37) is off. Hence $I_3 = I_4 = 0$. Also assume that only dc voltages are present and thus $I_1 = I_2 = 5\ \mu$A. Each of Q_7 to Q_{10} is biased at 10 µA, $V_{CC} = 3$ V, $V_{\text{BIAS3}} = V_{CC} - 1$, $R_7 = R_8 = 20$ kΩ, and $R_9 = R_{10} = 30$ kΩ. Neglect base currents and neglect the loading effect of the CMF circuit on the output nodes of the cascode circuit. The CMF circuit provides $V_B = V_{CM} + 0.4$.

(a) Determine the nominal values of V_B and V_{CM}. Does the value of V_{CM} ensure operation in the active mode for Q_7 through Q_{10}?
(b) If the CMF circuit were not present, what would be the change in v_{O1} and v_{O2} (i.e., in V_{CM}) as a result of a current mismatch $\Delta I = 0.3\ \mu A$ between Q_7-Q_8 and Q_9-Q_{10}? Use the output resistance values found in Example 13.7.
(c) Now, if the CMF circuit is connected, what change will it cause in V_B to eliminate the current mismatch ΔI? What is the corresponding change in V_{CM} from its nominal value?

Solution

(a) The nominal value of V_B is found as follows:

$$\begin{aligned} V_B &= V_{BE7} + (I_{E7} + I_1)R_7 \\ &\simeq 0.7 + (10+5) \times 10^{-3} \times 20 \\ &= 1\text{ V} \end{aligned}$$

The nominal value of V_{CM} can now be found from

$$V_{CM} = V_B - 0.4 = 1 - 0.4 = 0.6\text{ V}$$

For Q_7-Q_8 to be active,

$$V_{CM} > V_{B7,8} - 0.6$$

that is,

$$V_{CM} > 0.4\text{ V}$$

For Q_9-Q_{10} to be active

$$V_{CM} < V_{\text{BIAS3}} + 0.6$$

That is,

$$V_{CM} < V_{CC} - 1 + 0.6$$

resulting in

$$V_{CM} < 2.6\text{ V}$$

Thus, for all four cascode transistors to operate in the active mode,

$$0.4\text{ V} < V_{CM} < 2.6\text{ V}$$

Thus the nominal value of 0.6 V ensures active-mode operation.

Example 13.8 *continued*

(b) For $I_{C9} - I_{C7} = I_{C10} - I_{C8} = \Delta I$,

$$\Delta V_{CM} = \Delta I R_{o1}$$

where R_{o1} is the output resistance between the collectors of Q_7 and Q_9 and ground,

$$R_{o1} = R_{o7} \parallel R_{o9}$$

In Example 13.7 we found that $R_{o7} = 23\,\text{M}\Omega$ and $R_{o9} = 12.9\ \text{M}\Omega$; thus,

$$R_{o1} = 23 \parallel 12.9 = 8.3\,\text{M}\Omega$$

Thus,

$$\Delta V_{CM} = 0.3 \times 8.3 \simeq 2.5\ \text{V}$$

Now if ΔV_{CM} is positive,

$$V_{CM} = 0.6 + 2.5 = 3.1\ \text{V}$$

which exceeds the 2.6 V maximum allowed value before Q_9–Q_{10} saturate. If ΔV_{CM} is negative,

$$V_{CM} = 0.6 - 2.5 = -1.9\ \text{V}$$

which is far below the +0.4 V needed to keep Q_7–Q_8 in the active mode. Thus, in the absence of CMF, a current mismatch of ±0.3 μA would cause one set of the cascode transistors (depending on the polarity of ΔI) to saturate.

(c) With the CMF circuit in place, the feedback will adjust V_B by ΔV_B so that the currents in Q_7 and Q_8 will change by an increment equal to ΔI, thus restoring current equality. Since a change ΔV_B results in

$$\Delta I_{C7} = \Delta I_{C8} = \frac{\Delta V_B}{r_{e7} + R_7}$$

then

$$\Delta I = \frac{\Delta V_B}{r_{e7} + R_7}$$

$$\Delta V_B = \Delta I (r_{e7} + R_7)$$

$$= 0.3\ \mu\text{A}\left(\frac{25\,\text{mV}}{10\ \mu\text{A}} + 20\ \text{k}\Omega\right)$$

$$= 0.3 \times 22.5 = 6.75\,\text{mV}$$

Correspondingly

$$\Delta V_{CM} = \Delta V_B = 6.75\,\text{mV}$$

Thus, to restore the current equality, the change required in V_B and V_{CM} is only 6.75 mV.

13.4.5 Output-Stage Design for Near Rail-to-Rail Output Swing

As mentioned earlier, modern low-voltage bipolar op amps cannot afford to use the classical emitter-follower-based class AB output stage, which would consume too much of the power-supply voltage. Instead, a complementary pair of common-emitter transistors are utilized, as shown in Fig. 13.41. The output transistors Q_P and Q_N are operated in a class AB fashion. Typically, i_L can be as high as 10 mA to 15 mA and is determined by v_O and R_L. For $i_L = 0$, $i_P = i_N = I_Q$, where the quiescent current I_Q is normally a fraction of a milliamp.

The output stage in Fig. 13.41 is driven by two *separate but equal signals*, v_{BP} and v_{BN}. When v_{BP} and v_{BN} are high, Q_N supplies the load current in the direction opposite to that shown[9] and the output voltage v_O can swing to within 0.1 V or so of ground. In the meantime, Q_P is inactive. Nevertheless, in order to minimize crossover distortion, Q_P is prevented from turning off and is forced (as will be shown shortly) to conduct a minimum current of about $I_Q/2$.

The opposite happens when v_{BP} and v_{BN} are low: Q_P supplies the load current i_L in the direction indicated, and v_O can go up as high as $V_{CC} - 0.1$ V. In the meantime, Q_N is inactive but is prevented from turning off and forced to conduct a minimum current of about $I_Q/2$.

From the description above, we see that v_O can swing to within 0.1 V of each of the supply rails. This near rail-to-rail operation is the major advantage of this CE output stage. Its disadvantage is the relatively high output resistance. However, given that the op amp will almost always be used with a negative-feedback loop, the closed-loop output resistance can still be very low.

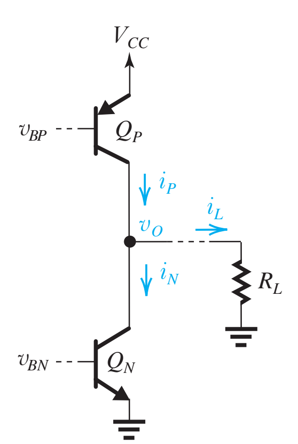


Figure 13.41 In order to provide v_O that can swing to within 0.1 V of V_{CC} and ground, a near rail-to-rail operation, the output stage utilizes common-emitter transistors. Note that the driving signals v_{BP} and v_{BN} are separate but identical.

A Buffer/Driver Stage The output transistors can be called on to supply currents in the 10 mA to 15 mA range. When this happens, the base currents of Q_P and Q_N can be substantial (recall that $\beta_P \simeq 10$ and $\beta_N \simeq 40$). Such large currents cannot usually be supplied directly by the amplifier stage preceding the output stage. Rather, a buffer/driver stage is usually needed, as shown in Fig. 13.42. Here an emitter follower Q_3 is used to drive Q_N. However, because of the low β_P, a double buffer consisting of complementary emitter followers Q_1 and Q_2 is used to drive Q_P. The driver stage is fed by two separate but identical signals v_{IP} and v_{IN} that come from the preceding amplifier stage (which is usually the second stage) in the op-amp circuit.[10]

[9] For this to happen, either R_L is returned to the positive supply (rather than ground) or R_L is capacitively coupled to the amplifier output.

[10] An interesting approach for generating two identical outputs in the second stage is utilized in the NE5234 (see Gray et al., 2009).

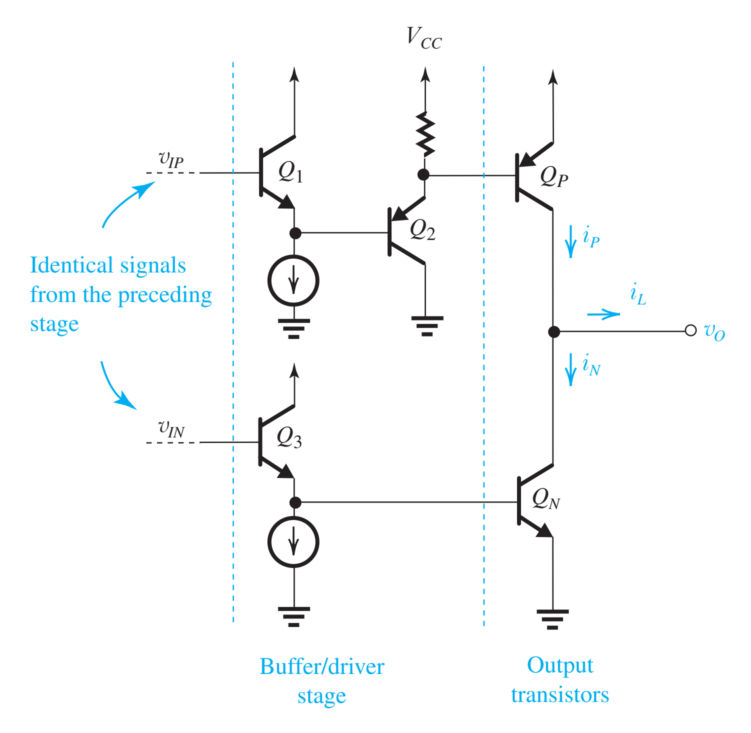


Figure 13.42 The output stage that is operated as class AB needs emitter-follower buffers/drivers to reduce the loading on the preceding stage and to provide the current gain necessary to drive Q_P and Q_N.

EXERCISE

13.39 (a) For the circuit in Fig. 13.42, find the current gain from each of the v_{IP} and v_{IN} terminals to the output in terms of β_P and β_N.

(b) For $i_L = \pm 10$ mA, how much signal current is needed at the v_{IP} and v_{IN} inputs?

Ans. (a) $\beta_N \beta_P^2$, β_N^2; (b) 2.5 μA, 6.25 μA

Establishing I_Q and Maintaining a Minimum Current in the Inactive Transistor

We next consider the circuit for establishing the quiescent current I_Q in Q_N and Q_P and for maintaining a minimum current of $I_Q/2$ in the inactive output transistor. Figure 13.43 shows a fuller version of the output stage. In addition to the output transistors $Q_P - Q_N$ and the buffer/driver stage, which we have already discussed, the circuit includes two circuit blocks whose operation we shall now explain.

The first is the circuit composed of the differential pair $Q_6 - Q_7$ and associated transistors Q_4 and Q_5, and resistors R_4 and R_5. This circuit measures the currents in the output transistors,

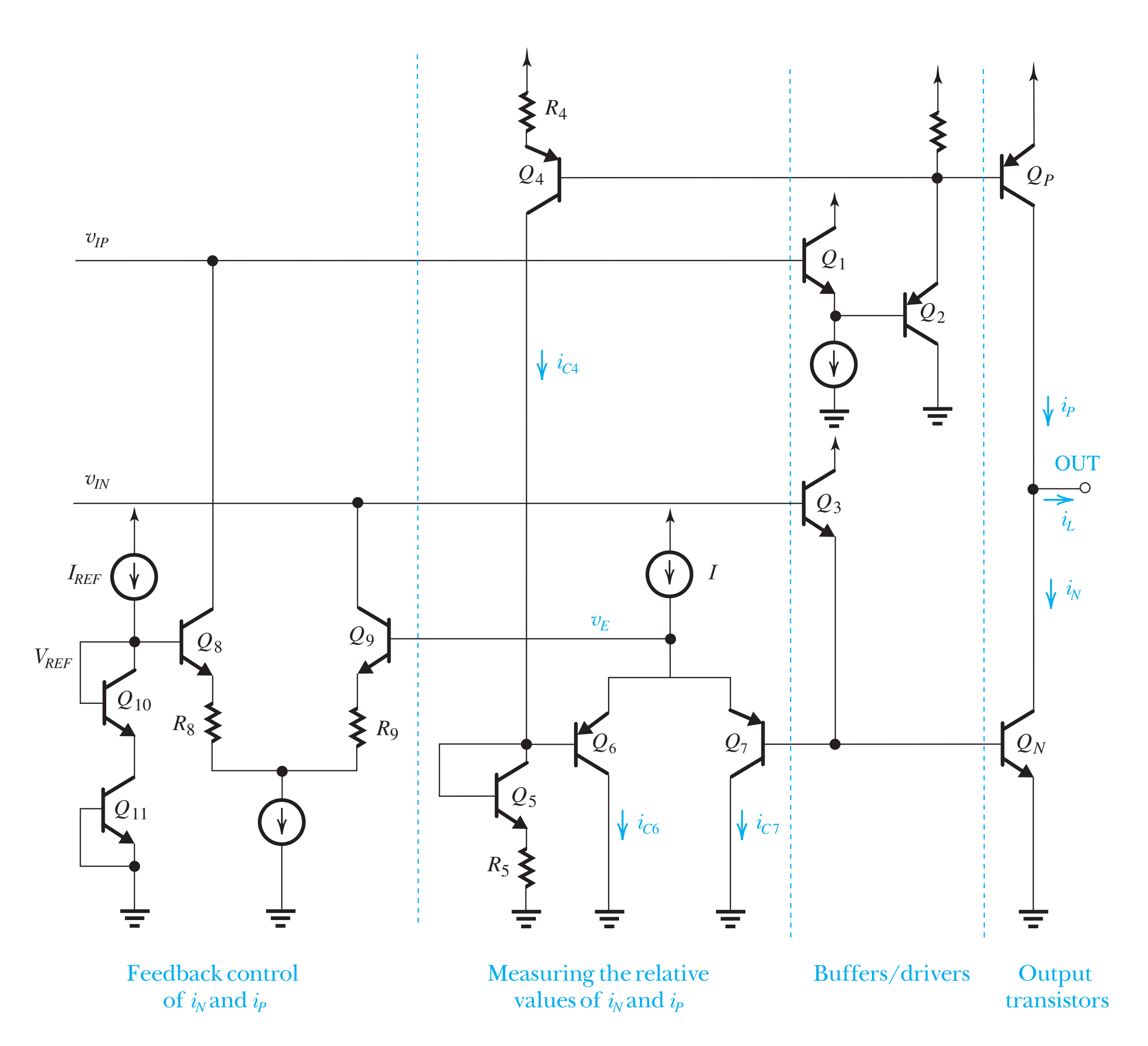


Figure 13.43 A more complete version of the output stage showing the circuits that establish the quiescent current in Q_P and Q_N. As well, this circuit forces a minimum current of $(I_Q/2)$ to follow in the inactive output transistor, preventing the transistor from turning off and thus minimizing crossover distortion.

i_P and i_N, and arranges for the current I to divide between Q_6 and Q_7 according to the ratio i_N/i_P, and provides a related output voltage v_E. Specifically, it can be shown (Problems 13.86 and 13.87) that

$$i_{C6} = I\,\frac{i_N}{i_P + i_N} \tag{13.130}$$

$$i_{C7} = I\,\frac{i_P}{i_P + i_N} \tag{13.131}$$

$$v_E = V_T \ln\left[\frac{i_N i_P}{i_N + i_P}\,\frac{I}{I_{SN} I_{S7}}\right] \tag{13.132}$$

where I_{SN} and I_{S7} are the saturation currents of Q_N and Q_7, respectively. Observe that for $i_P \gg i_N$, $i_{C6} \simeq 0$ and $i_{C7} \simeq I$. Thus Q_6 turns off and Q_7 conducts all of I. The emitter voltage v_E becomes

$$v_E \simeq V_T \ln\left(\frac{i_N}{I_{SN}}\right) + V_T \ln\left(\frac{I}{I_{S7}}\right)$$

Thus,

$$v_E = V_T \ln\left(\frac{i_N}{I_{SN}}\right) + V_{EB7} \tag{13.133}$$

This equation simply states that $v_E = v_{BEN} + V_{EB7}$, which could have been directly obtained from the circuit diagram in Fig. 13.43. The important point to note, however, is that since V_{EB7} is a constant, v_E is determined by the current i_N in the inactive transistor, Q_N. In the other extreme case of $i_N \gg i_P$, $i_{C6} \simeq I$, $i_{C7} \simeq 0$; thus Q_7 turns off and Q_6 conducts all of I. In this case we can use Eq. (13.132) to show that

$$v_E = V_T \ln\left(\frac{i_P}{I_{SN}}\right) + V_{EB6} \tag{13.134}$$

Thus, here too, since V_{EB6} is a constant, v_E is determined by the current in the inactive transistor, Q_P.

The second circuit block is a differential amplifier composed of Q_8–Q_9 with their emitter-degeneration resistors R_8, R_9. The voltage v_E generated by the measuring circuit is fed to one input of the differential amplifier, and the other input is fed with a reference voltage V_{REF} generated by passing a reference current I_{REF} through the series connection of diode-connected transistors Q_{10} and Q_{11}. This differential amplifier takes part in a negative-feedback loop that uses the value of v_E to control the currents i_P and i_N through the nodes v_{IP} and v_{IN}. The objective of the feedback control is to set the current in the inactive output transistor to a minimum value. To see how the feedback operates, consider the case when $i_P \gg i_N$, and thus Q_N is the inactive transistor. In this case, Q_6 turns off, Q_7 conducts all of I, and v_E is given by Eq. (13.133). Now, if for some reason i_N falls below its minimum intended value, v_E decreases, causing i_{C9} to decrease. This in turn will cause the node v_{IN} to rise and the voltage at the base of Q_N will eventually rise, thus increasing i_N to its intended value.

Analytically, we can obtain a relationship between i_N and i_P as follows. Assume that the loop gain of the feedback loop that is anchored by the differential amplifier Q_8–Q_9 is high enough to force the two input terminals to the same voltage, that is,

$$v_E = V_{\text{REF}} = V_T \ln \frac{I_{\text{REF}}}{I_{S10}} + V_T \ln \frac{I_{\text{REF}}}{I_{S11}}$$

Substituting for v_E from Eq. (13.132) results in

$$\frac{i_N i_P}{i_N + i_P} = \left(\frac{I_{\text{REF}}^2}{I}\right)\left(\frac{I_{SN}}{I_{S10}}\right)\left(\frac{I_{S7}}{I_{S11}}\right) \tag{13.135}$$

Observe that the quantity on the right-hand side is a constant. In the quiescent case, $i_N = i_P = I_Q$, Eq. (13.135) yields

$$I_Q = 2\left(\frac{I_{\text{REF}}^2}{I}\right)\left(\frac{I_{SN}}{I_{S10}}\right)\left(\frac{I_{S7}}{I_{S11}}\right) \tag{13.136}$$

Thus, the constant on the right-hand side of Eq. (13.135) is $I_Q/2$, and we can rewrite (13.135) as

$$\frac{i_N i_P}{i_N + i_P} = \frac{1}{2} I_Q \tag{13.137}$$

Equation (13.137) clearly shows that for $i_N \gg i_P$, $i_P \simeq \frac{1}{2}I_Q$, and that for $i_P \gg i_N$, $i_N \simeq \frac{1}{2}I_Q$. Thus the circuit not only establishes the quiescent current I_Q (Eq. 13.136) but also sets the minimum current in the inactive output transistor at $\frac{1}{2}I_Q$.

EXERCISE

D13.40 For the circuit in Fig. 13.43, determine the value that I_{REF} should have so that Q_N and Q_P have a quiescent current $I_Q = 0.4$ mA. Assume that the transistor areas are scaled so that $I_{SN}/I_{S10} = 10$ and $I_{S7}/I_{S11} = 2$. Let $I = 10\ \mu$A. Also, if i_L in the direction out of the amplifier is 10 mA, find i_P and i_N.
Ans. $I_{REF} = 10\ \mu$A; $i_P \simeq 10.2$ mA, $i_N \simeq 0.2$ mA

13.4.6 Concluding Remark

This section presented a sample of the design techniques employed in modern BJT op amps.

Summary

- Most CMOS op amps are designed to operate as part of a VLSI circuit and thus are required to drive only small capacitive loads. Therefore, most do not have a low-output-resistance stage.
- There are basically two approaches to the design of CMOS op amps: a two-stage configuration and a single-stage topology utilizing the folded-cascode circuit.
- In the two-stage CMOS op amp, approximately equal gains are realized in the two stages.
- The threshold mismatch ΔV_t together with the low transconductance of the input stage result in a larger input offset voltage for CMOS op amps than for bipolar units.
- Miller compensation is employed in the two-stage CMOS op amp, but a series resistor is required to place the transmission zero at either $s = \infty$ or on the negative real axis so as to increase the phase margin.
- CMOS op amps have higher slew rates than their bipolar counterparts with comparable f_t values.
- Use of the cascode configuration increases the gain of a CMOS amplifier stage by about two orders of magnitude, thus making possible a single-stage op amp.
- The dominant pole of the folded-cascode op amp is determined by the total capacitance at the output node, C_L. Increasing C_L improves the phase margin at the expense of reducing the bandwidth.
- By using two complementary input differential pairs in parallel, the input common-mode range can be extended to equal the entire power-supply voltage, providing so-called rail-to-rail operation at the input.
- The output voltage swing of the folded-cascode op amp can be extended by utilizing a wide-swing current mirror in place of the cascode mirror.
- The internal circuit of the 741 op amp embodies many of the design techniques employed in bipolar analog integrated circuits.
- The 741 circuit consists of an input differential stage, a high-gain single-ended second stage, and a class AB output stage. Though more than 40 years old, this structure is typical of most BJT op amps and is known as the two-stage topology (not counting the output stage). It is also the same structure used in the two-stage CMOS op amp of Section 13.1.

- To obtain low input offset voltage and current, and high CMRR, the 741 input stage is designed to be perfectly balanced. The CMRR is increased by common-mode feedback, which also stabilizes the dc operating point.
- To obtain high input resistance and low input bias current, the input stage of the 741 is operated at a very low current level.
- In the 741, output short-circuit protection is accomplished by turning on a transistor that takes away most of the base current drive of the output transistor.
- The use of Miller frequency compensation in the 741 circuit enables locating the dominant pole at a very low frequency while utilizing a relatively small compensating capacitance.
- Two-stage op amps can be modeled as a transconductance amplifier feeding an ideal integrator with C_C as the integrating capacitor.
- The slew rate of a two-stage op amp is determined by the first-stage bias current and the frequency-compensation capacitor.
- While the 741 and its generation of op amps nominally operate from ±15-V power supplies, modern BJT op amps typically utilize a single ground-referenced supply of only 2 V to 3 V.
- Operation from a single low-voltage supply gives rise to a number of new important specifications including a common-mode input range that extends beyond the supply rails (i.e., more than rail-to-rail operation) and a near rail-to-rail output voltage swing.
- The rail-to-rail input common-mode range is achieved by using resistive loads (instead of current-mirror loads) for the input differential pair as well as utilizing two complementary differential amplifiers in parallel.
- To increase the gain of the input stage above that achieved with resistive loads, the folded-cascode configuration is utilized.
- To regulate the dc bias voltages at the outputs of the differential folded-cascode stage so as to maintain active-mode operation at all times, common-mode feedback is employed.
- The output stage of a low-voltage op amp utilizes a complementary pair of common-emitter transistors. This allows v_O to swing to within 0.1 V or so from each of the supply rails. The disadvantage is a high open-loop output resistance. This, however, is substantially reduced when negative feedback is applied around the op amp.
- Modern output stages operate in the class AB mode and utilize interesting feedback techniques to set the quiescent current as well as to ensure that the inactive output transistor does not turn off, a precaution that avoids increases in crossover distortion.

PROBLEMS

Computer Simulation Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as allowable signal swing and amplifier nonlinear distortion. Instructions to assist in setting up PSpice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 13.1: The Two-Stage CMOS Op Amp

13.1 A particular design of the two-stage CMOS operational amplifier of Fig. 13.1 utilizes ±1-V power supplies. All transistors are operated at overdrive voltages of 0.2-V magnitude. The process technology provides devices with $V_{tn} = |V_{tp}| = 0.4$ V. Find the input common-mode range and the range allowed for v_O.

13.2 The CMOS op amp of Fig. 13.1 is fabricated in a process for which $V'_{An} = 25$ V/μm and $|V'_{Ap}| = 20$ V/μm. Find A_1, A_2,

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

Transistor	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8
W/L (μm/μm)	36/0.3	36/0.3	6/0.3	6/0.3	30/0.3	W/0.3	45/0.3	6/0.3

and A_v if all devices are 0.3 μm long, Q_1 and Q_2 are operated at overdrive voltages of 0.15-V magnitude, and Q_6 is operated at $V_{OV} = 0.2$ V. Also, determine the op-amp output resistance obtained when the second stage is biased at 0.3 mA. What do you expect the output resistance of a unity-gain voltage amplifier to be, using this op amp?

D 13.3 The CMOS op amp of Fig. 13.1 is fabricated in a process for which $|V_A'|$ for all devices is 20 V/μm. If all transistors have $L = 0.3$ μm and are operated at equal overdrive voltages, find the magnitude of the overdrive voltage required to obtain a dc open-loop gain of 1600 V/V.

13.4 Consider the circuit in Fig. 13.1 with the device geometries shown at the top of this page. Let $I_{\text{REF}} = 40$ μA, $|V_t|$ for all devices $= 0.45$ V, $\mu_n C_{ox} = 270$ μA/V^2, $\mu_p C_{ox} = 70$ μA/V^2, $|V_A|$ for all devices $= 15$ V, $V_{DD} = V_{SS} = 1$ V. Determine the width of Q_6, W, that will ensure that the op amp will not have a systematic offset voltage. Then, for all devices, evaluate I_D, $|V_{OV}|$, $|V_{GS}|$, g_m, and r_o. Provide your results in a table. Also find A_1, A_2, the dc open-loop voltage gain, the input common-mode range, and the output voltage range. Neglect the effect of V_A on the bias currents.

D 13.5 Design the two-stage CMOS op amp in Fig. 13.1 to provide a CMRR of about 72 dB. If all the transistors are operated at equal overdrive voltages of 0.15 V and have equal channel lengths, find the minimum required channel length. For this technology, $|V_A'| = 15$ V/μm. What is the dc gain realized?

13.6 A two-stage CMOS op amp has $G_{m1} = 0.8$ mA/V, $G_{m2} = 2.4$ mA/V, $C_1 = 0.1$ pF, and $C_2 = 1.2$ pF. Find the value of C_C that will provide a unity-gain frequency of 120 MHz. Also, determine the values of f_{P2} and f_Z.

13.7 For the CMOS amplifier in Fig. 13.1, whose equivalent circuit is shown in Fig. 13.2, let $G_{m1} = 1$ mA/V, $R_1 = 100$ kΩ, $C_1 = 0.1$ pF, $G_{m2} = 2$ mA/V, $R_2 = 50$ kΩ, and $C_2 = 2$ pF.

(a) Find the dc gain.
(b) Without C_C connected, find the frequencies of the two poles in radians per second and sketch a Bode plot for the gain magnitude.
(c) With C_C connected, find ω_{P2}. Then find the value of C_C that will result in a unity-gain frequency ω_t at least two octaves below ω_{P2}. For this value of C_C, find ω_{P1} and ω_Z and sketch a Bode plot for the gain magnitude.

13.8 A CMOS op amp with the topology in Fig. 13.1 has $g_{m1} = g_{m2} = 1$ mA/V, $g_{m6} = 3$ mA/V, the total capacitance between node D_2 and ground is 0.2 pF, and the total capacitance between the output node and ground is 3 pF. Find the value of C_C that results in $f_t = 50$ MHz, and verify that f_t is lower than f_Z and f_{P2}.

13.9 A particular design of the two-stage CMOS op amp of Fig. 13.1 has $G_{m1} = 1$ mA/V and $G_{m2} = 2$ mA/V. The total capacitance at the output node is 1 pF. While utilizing a Miller compensation capacitor C_C without a series resistance R, the amplifier is made to have a uniform −20-dB/decade gain rolloff with a unity-gain frequency f_t of 100 MHz.

(a) What must the value of C_C be?
(b) What do you estimate the frequencies of the poles, f_{P1} and f_{P2}, and of the right-half-plane zero, f_Z, to be?
(c) What is the phase margin obtained?
(d) To increase the phase margin, a resistance R is connected in series with C_C. What is the value of R that results in $f_Z = \infty$, and what is the resulting phase margin?
(e) If R is increased further, until it moves the zero into the left half-plane and thus turns the phase it introduces into phase lead, what value of R is needed to obtain a phase margin of 85°?

D 13.10 A particular implementation of the CMOS amplifier of Figs. 13.1 and 13.2 provides $G_{m1} = 0.3$ mA/V,

$G_{m2} = 0.6$ mA/V, $r_{o2} = r_{o4} = 222$ kΩ, $r_{o6} = r_{o7} = 111$ kΩ, and $C_2 = 1$ pF.

(a) Find the dc gain.
(b) Find the frequency of the second pole, f_{P2}.
(c) Find the value of the resistance R that, when placed in series with C_C, causes the transmission zero to be located at $s = \infty$.
(d) With R in place, as in (c), find the value of C_C that results in the highest possible value of f_t while providing a phase margin of 80°. What value of f_t is realized? What is the corresponding frequency of the dominant pole?
(e) To what value should C_C be changed to double the value of f_t? At the new value of f_t, what is the phase shift introduced by the second pole? To reduce this excess phase shift to 10° and thus obtain an 80° phase margin, as before, what value should R be changed to?

13.11 A two-stage CMOS op amp has each of its first-stage transistors Q_1 and Q_2 operating at an overdrive voltage of 0.2 V. The op amp has a uniform −20-dB/decade frequency response with a unity-gain frequency of 100 MHz. What do you expect the slew rate of this amplifier to be? If each of Q_1 and Q_2 is biased at 50 μA, what must the value of C_C be?

D 13.12 A two-stage CMOS op amp similar to that in Fig. 13.1 is found to have a capacitance between the output node and ground of 0.7 pF. If it is desired to have a unity-gain bandwidth f_t of 100 MHz with a phase margin of 72° what must g_{m6} be set to? Assume that a resistance R is connected in series with the frequency-compensation capacitor C_C and adjusted to place the transmission zero at infinity. What value should R have? If the first stage is operated at $|V_{OV}| = 0.15$ V, what is the value of slew rate obtained? If the first-stage bias current $I = 100$ μA, what is the required value of C_C?

D 13.13 A CMOS op amp with the topology shown in Fig. 13.1 is designed to provide $G_{m1} = 1$ mA/V and $G_{m2} = 5$ mA.

(a) Find the value of C_C that results in $f_t = 80$ MHz.
(b) What is the maximum value that C_2 can have while achieving a 70° phase margin?

D 13.14 A CMOS op amp with the topology shown in Fig. 13.1 but with a resistance R included in series with C_C is designed to provide $G_{m1} = 0.8$ mA/V and $G_{m2} = 2$ mA/V.

(a) Find the value of C_C that results in $f_t = 100$ MHz.
(b) For $R = 500$ Ω, what is the maximum allowed value of C_2 for which a phase margin of at least 60° is obtained?

13.15 A two-stage CMOS op amp resembling that in Fig. 13.1 is found to have a slew rate of 60 V/μs and a unity-gain bandwidth f_t of 60 MHz.

(a) Estimate the value of the overdrive voltage at which the input-stage transistors are operating.
(b) If the first-stage bias current $I = 120$ μA, what value of C_C must be used?
(c) For a process for which $\mu_p C_{ox} = 60$ μA/V^2, what W/L ratio applies for Q_1 and Q_2?

D 13.16 Sketch the circuit of a two-stage CMOS amplifier having the structure of Fig. 13.1 but utilizing NMOS transistors in the input stage (i.e., Q_1 and Q_2).

D 13.17 (a) Show that the PSRR$^-$ of a CMOS two-stage op amp for which all transistors have the same channel length and are operated at equal $|V_{OV}|$ is given by

$$\text{PSRR}^- = 2\left|\frac{V_A}{V_{OV}}\right|^2$$

(b) For $|V_{OV}| = 0.15$ V, what is the minimum channel length required to obtain a PSRR$^-$ of 72 dB? For the technology available, $|V_A'| = 15$ V/μm.

D 13.18 It is required to design the circuit of Fig. 13.8 to provide a bias current I_{REF} of 225 μA with Q_8 and Q_9 as matched devices having $W/L = 60/0.5$. Transistors Q_{10}, Q_{11}, and Q_{13} are to be identical and must have the same g_m as Q_8 and Q_9. Transistor Q_{12} is to be four times as wide as Q_{13}. Let $k_n' = 3k_p' = 180$ μA/V^2, and $V_{tn} = |V_{tp}| = 0.5$ V; let all channel lengths be equal; and let $V_{DD} = V_{SS} = 1.5$ V. Find the required value of R_B. What is the voltage drop across R_B? Also specify the W/L ratios of Q_{10}, Q_{11}, Q_{12}, and Q_{13} and give the expected dc voltages at the gates of Q_{12}, Q_{10}, and Q_8.

Section 13.2: The Folded-Cascode CMOS Op Amp

D 13.19 The op-amp circuit of Fig. 13.10 is operated from ±1-V power supplies. If the power dissipated in the circuit is to be limited to 1 mW, find the maximum value of I_B allowed. If this value is used, and each of Q_1 and Q_2 is to be biased at a current four times that used for each of Q_3 and Q_4, find the value of I, $I_{D1,2}$, and $I_{D3,4}$.

D 13.20 For the folded-cascode op amp in Fig. 13.10 utilizing power supplies of ±1 V, find the values of V_{BIAS1}, V_{BIAS2}, and V_{BIAS3} to maximize the allowable range of V_{ICM} and v_O. Assume that all transistors are operated at equal overdrive voltages of 0.15 V. Assume $|V_t|$ for all devices is 0.4 V. Specify the maximum range of V_{ICM} and of v_O.

D 13.21 For the folded-cascode op-amp circuit of Figs. 13.9 and 13.10 with bias currents $I = 400\ \mu\text{A}$ and $I_B = 250\ \mu\text{A}$, and with all transistors operated at overdrive voltages of 0.2 V, find the W/L ratios for all devices. Assume that the technology available is characterized by $k_n' = 400\ \mu\text{A/V}^2$ and $k_p' = 100\ \mu\text{A/V}^2$.

13.22 Consider a design of the cascode op amp of Fig. 13.10 for which $I = 400\ \mu\text{A}$ and $I_B = 250\ \mu\text{A}$. Assume that all transistors are operated at $|V_{OV}| = 0.2$ V and that for all devices, $|V_A| = 10$ V. Find G_m, R_o, and A_v. Also, if the op amp is connected in the feedback configuration shown in Fig. P13.22, find the voltage gain and output resistance of the closed-loop amplifier.

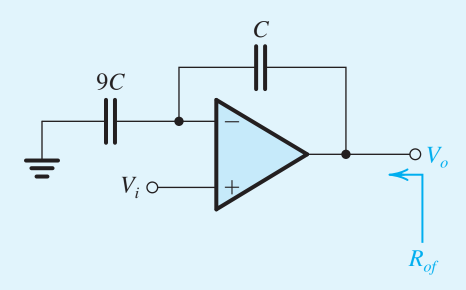


Figure P13.22

D 13.23 Consider the folded-cascode op amp of Fig. 13.9 when loaded with a 10-pF capacitance. What should the bias current I_B be to obtain a slew rate of at least 10 V/μs? If the input-stage transistors are biased at a current three times that at which each of Q_3 and Q_4 is biased, find the value of I. If the input-stage transistors are operated at overdrive voltages of 0.15 V, what is the unity-gain bandwidth realized? If the two nondominant poles have the same frequency of 50 MHz, what is the phase margin obtained? If it is required to have a phase margin of 75°, what must f_t be reduced to? By what amount should C_L be increased? What is the new value of SR?

13.24 For a particular design of the folded-cascode op amp in Fig. 13.9, $I < I_B$. What slew rate is obtained?

D *13.25 Design the folded-cascode circuit of Fig. 13.10 to provide voltage gain of 80 dB and a unity-gain frequency of 20 MHz when $C_L = 10$ pF. Design for $I_B = I$, and operate all devices at the same $|V_{OV}|$. Utilize transistors with 1-μm channel length for which $|V_A|$ is specified to be 12 V. Find the required overdrive voltages and bias currents. What slew rate is achieved? Also, for $k_n' = 2.5k_p' = 400\ \mu\text{A/V}^2$, specify the required width of each of the 11 transistors used.

D 13.26 Sketch the circuit that is complementary to that in Fig. 13.10, that is, one that uses an input p-channel differential pair.

****13.27** This problem presents a very interesting addition to the folded-cascode op-amp circuit of Fig. 13.10, designed to deal with the situation during amplifier slewing. In particular, the additional circuitry does two things: It prevents Q_1 and Q_{11} from going into the triode region, and it increases the current available to charge C_L and thus increases the slew rate. The circuit is shown in Fig. P13.27 (with the current-mirror circuit omitted, for simplicity). Observe that three transistors are added: Q_{14}, which is biased by a constant-current source (20 μA), establishes the dc currents in Q_9 and Q_{10}. Assume with respect to Q_9 and Q_{10} that each has a W/L ratio 10 times that of Q_{14}. The other two additional transistors are Q_{12} and Q_{13}, which are diode connected and are normally cut off.

(a) For $V_{id} = 0$, find the bias current in each of Q_1, Q_2, Q_3, Q_4, Q_{14}, Q_9, and Q_{10}. Also, for the dc voltages shown, and assuming $V_{tn} = |V_{tp}| = 0.45$ V and that all conducting devices are operating at $|V_{OV}| = 0.15$ V, show that Q_{12} and Q_{13} will be cut off.

(b) For an input differential signal that causes Q_2 to turn off and Q_1 to conduct the entire bias current (320 μA), Q_{12} turns on (while Q_{13} remains off). Noting that the drain current of Q_{12} adds to the 20 μA flowing through Q_{14}, find the current that now flows through Q_{10} and onto C_L.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

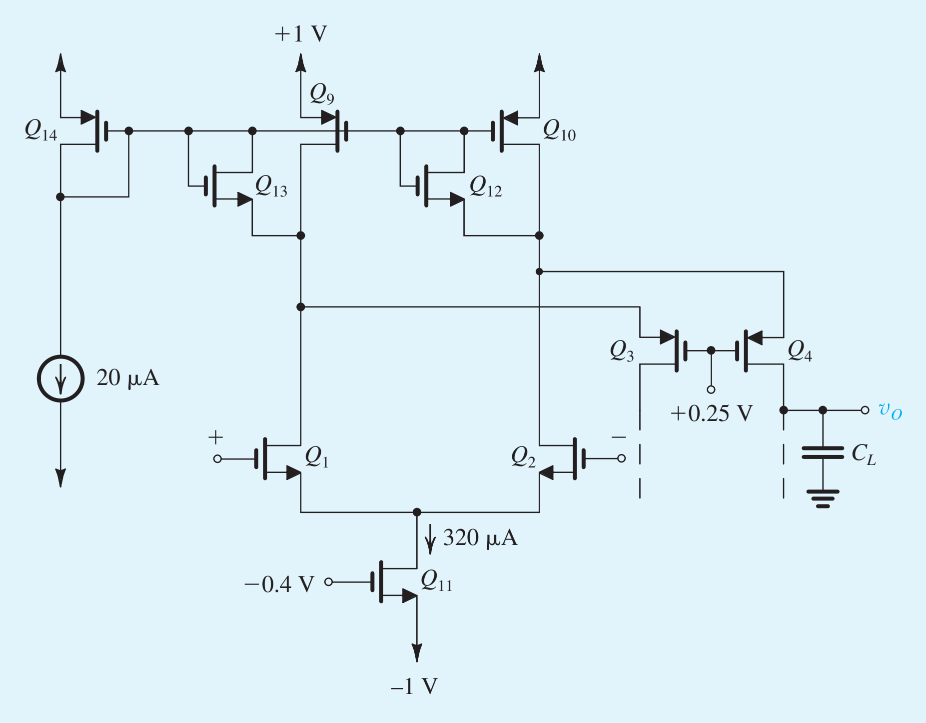


Figure P13.27

By what factor does the slew rate increase relative to the value without Q_{12} present? Also give an approximate estimate of the voltage at the drain of Q_1 during the slewing transient.

13.28 For the circuit in Fig. 13.12, assume that all transistors are operating at equal overdrive voltages of 0.15-V magnitude and have $|V_t| = 0.45$ V and that $V_{DD} = V_{SS} = 1$ V. Find (a) the range over which the NMOS input stage operates, (b) the range over which the PMOS input stage operates, (c) the range over which both operate (the overlap range), and (d) the input common-mode range. Assume that all current sources require a minimum voltage of $|V_{OV}|$ to operate properly.

13.29 A particular design of the wide-swing current mirror of Fig. 13.13(b) utilizes devices having $W/L = 20$, $k_n' = 400\ \mu\text{A/V}^2$, and $V_t = 0.45$ V. For $I_{\text{REF}} = 90\ \mu$A, what value of V_{BIAS} is needed? Also give the voltages that you expect to appear at all nodes and specify the minimum voltage allowable at the output terminal. If V_A is specified to be 10 V, what is the output resistance of the mirror?

D 13.30 For the folded-cascode circuit of Fig. 13.9, let the total capacitance to ground at each of the source nodes of Q_3 and Q_4 be denoted C_P. Assuming that the incremental resistance between the drain of Q_3 and ground is small, show that the pole that arises at the interface between the first and second stages has a frequency $f_P \simeq g_{m3}/2\pi C_P$. Now, if this is the only nondominant pole, what is the largest value that C_P can be (expressed as a fraction of C_L) while a phase margin of 80° is achieved? Assume that all transistors are operated at the same bias current and overdrive voltage.

Section 13.3: The 741 BJT Op Amp

13.31 In the 741 op-amp circuit of Fig. 13.14, Q_1, Q_2, Q_5, and Q_6 are biased at collector currents of 9.5 μA; Q_{16} is biased at a collector current of 16.2 μA; and Q_{17} is biased at a collector current of 550 μA. All these devices are of the “standard *npn*” type, having $I_S = 10^{-14}$ A, $\beta = 200$, and $V_A = 125$ V. For each of these transistors, find V_{BE}, g_m, r_e, r_π, and r_o. Provide your results in table form. (Note that these parameter values are utilized in the text in the analysis of the 741 circuit.)

D 13.32 For the circuit in Fig. P13.32, neglect base currents and use the exponential i_C–v_{BE} relationship to show that

$$I_3 = I_1\sqrt{\frac{I_{S3}I_{S4}}{I_{S1}I_{S2}}}$$

Find I_1 for the case in which $I_{S3} = I_{S4} = 3 \times 10^{-14}$ A, $I_{S1} = I_{S2} = 10^{-14}$ A, and a bias current $I_3 = 150$ μA is required.

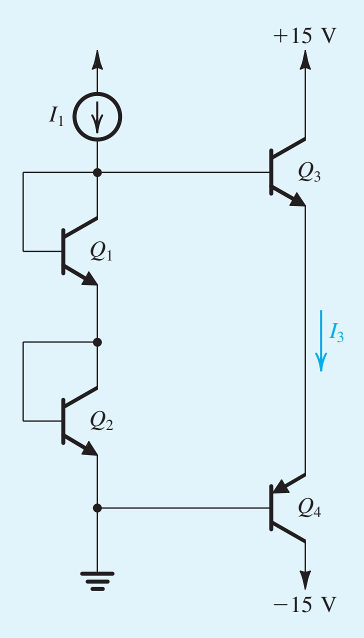


Figure P13.32

13.33 Transistor Q_{13} in the circuit of Fig. 13.14 consists, in effect, of two transistors whose emitter–base junctions are connected in parallel and for which $I_{SA} = 0.25 \times 10^{-14}$ A, $I_{SB} = 0.75 \times 10^{-14}$ A, $\beta = 50$, and $V_A = 50$ V. For operation at a total emitter current of 0.73 mA, find values for the parameters V_{EB}, g_m, r_e, r_π, and r_o for the A and B devices.

13.34 In the circuit of Fig. 13.14, Q_1 and Q_2 exhibit emitter–base breakdown at 7 V, while for Q_3 and Q_4 such a breakdown occurs at about 50 V. What differential input voltage would result in the breakdown of the input-stage transistors?

D 13.35 Figure P13.35 shows the CMOS version of the circuit in Fig. P13.32. Find the relationship between I_3 and I_1 in terms of k_1, k_2, k_3, and k_4 of the four transistors, assuming the threshold voltages of all devices to be equal in magnitude. Note that k denotes $\mu C_{ox}W/L$. In the event that $k_1 = k_2$ and $k_3 = k_4 = 16k_1$, find the required value of I_1 to yield a bias current in Q_3 and Q_4 of 1.6 mA.

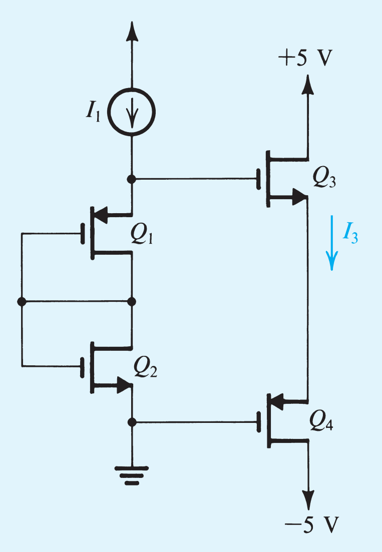


Figure P13.35

D 13.36 For the 741 circuit, estimate the reference current I_{REF} in the event that ±10-V supplies are used. What value of R_5 would be necessary to reestablish the same bias current for ±10-V supplies as exists for ±15 V in the original design?

D 13.37 Design a Widlar current source to supply a current of 10 μA given a reference input current of 0.3 mA. Assume that the transistors have $I_S = 10^{-14}$ A and high β. Find V_{BE} of each of the two transistors as well as the value of R.

13.38 Consider the dc analysis of the 741 input stage shown in Fig. 13.15.

(a) Derive an expression for I taking β_P into account. What is the percentage change in I if β_P drops from 50 to 20?
(b) Now, consider an alternative design of this circuit in which the feedback loop is eliminated. That is, Q_8 and Q_9 are eliminated and I_{C10} is fed to the common-base connection of Q_3 and Q_4. What is I now in terms of I_{C10}? If β_P changes from 50 to 20, what is the resulting percentage change in I?

D 13.39 Consider the dc analysis of the 741 input stage shown in Fig. 13.15 for the situation in which $I_{S9} = 2I_{S8}$. For $I_{C10} = 19$ μA and assuming β_P to be high, what does I become? Redesign the Widlar source to reestablish $I_{C1} = I_{C2} = 9.5$ μA.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

D 13.40 Consider the circuit shown in Fig. 13.15. If $I_{C10} =$ 40 μA and I is required to be 10 μA, what must be the ratio of the emitter–junction area of Q_9 to that area of Q_8? Assume that β_P is large.

13.41 For the mirror circuit shown in Fig. 13.16 with the bias and component values given in the text for the 741 circuit, what does the current in Q_6 become if R_2 is shorted?

D 13.42 It is required to redesign the circuit of Fig. 13.16 by selecting a new value for R_3 so that when the base currents are *not* neglected, the collector currents of Q_5, Q_6, and Q_7 all become equal, assuming that the input current $I_{C3} = 9.5$ μA. Find the new value of R_3 and the three currents. Recall that $\beta_N = 200$.

13.43 Consider the input circuit of the 741 op amp of Fig. 13.14 when the emitter current of Q_8 is about 19 μA. If β of Q_1 is 150 and that of Q_2 is 220, find the input bias current I_B and the input offset current I_{OS} of the op amp.

13.44 For a particular application, consideration is being given to selecting 741 ICs for input bias and offset currents limited to 60 nA and 5 nA, respectively. Assuming other aspects of the selected units to be normal, what minimum β_N and what β_N variation are implied?

13.45 For a 741 employing ±5-V supplies, $|V_{BE}| = 0.6$ V, and $|V_{CEsat}| = 0.2$ V, find the input common-mode range. Neglect the voltage drops across R_1 and R_2.

D 13.46 Consider the design of the second stage of the 741. What value of R_9 would be needed to reduce I_{C16} to 9.5 μA? (*Hint:* Build on Exercise 13.21)

D 13.47 Reconsider the 741 output stage as shown in Fig. 13.17, in which R_{10} is adjusted to make $I_{C19} = I_{C18}$. What is the new value of R_{10}? What values of I_{C14} and I_{C20} result? Recall that $I_{REF} = 0.73$ mA.

D *13.48 An alternative approach to providing the voltage drop needed to bias the output transistors is the V_{BE}–multiplier circuit shown in Fig. P13.48. Design the circuit to provide a terminal voltage of 1.118 V (the same as in the 741 circuit). Base your design on half the current flowing through R_1, and assume that $I_S = 10^{-14}$ A and $\beta = 200$. What is the incremental resistance between the two terminals of the V_{BE}–multiplier circuit?

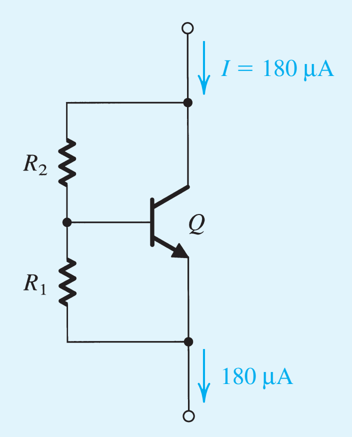


Figure P13.48

13.49 For the circuit of Fig. 13.14, what is the total current required from the power supplies when the op amp is operated in the linear mode, but with no load? Hence, estimate the quiescent power dissipation in the circuit. (*Hint:* Use the data given in Table 13.1.)

13.50 Consider the 741 input stage as modeled in Fig. 13.18, with two additional *npn* diode-connected transistors, Q_{1a} and Q_{2a}, connected between the present *npn* and *pnp* devices, one per side. Convince yourself that each of the additional devices will be biased at the same current as Q_1 to Q_4—that is, 9.5 μA. What does R_{id} become? What does G_{m1} become? What is the value of R_{o4} now? What is the output resistance of the first stage, R_{o1}? (Note that R_{o6} remains unchanged at 18.2 MΩ.) What is the new open-circuit voltage gain, $G_{m1}R_{o1}$? Compare these values with the original ones, namely, $R_{id} = 21$ MΩ, $G_{m1} = 0.19$ mA/V, $R_{o4} = 10.5$ MΩ, $R_{o1} = 6.7$ MΩ, and $|A_{vo}| = 1273$ V/V.

13.51 Consider the current mirror in Fig. 13.19. What value must R_2 be increased to in order to increase R_{o6} by a factor of 2? Recall that Q_6 is operating at $I_{C6} = 9.5$ μA and has $\beta = 200$ and $V_A = 125$ V.

13.52 Repeat Exercise 13.24 with $R_1 = R_2$ replaced by 2-kΩ resistors.

13.53 A manufacturing problem in a 741 op amp causes the current-transfer ratio of the mirror circuit that loads the input stage to become 0.8 A/A. For input devices (Q_1–Q_4) appropriately matched and with high β, and normally biased at 9.5 μA, what input offset voltage results?

***13.54** In Example 13.4 we investigated the effect of a mismatch between R_1 and R_2 on the input offset voltage of the op amp. Conversely, R_1 and R_2 can be deliberately mismatched (using the circuit shown in Fig. P13.54, for example) to compensate for the op-amp input offset voltage.

(a) Show that an input offset voltage V_{OS} can be compensated for (i.e., reduced to zero) by creating a relative mismatch $\Delta R/R$ between R_1 and R_2,

$$\frac{\Delta R}{R} = \frac{V_{OS}}{2V_T}\frac{1+r_e/R}{1-V_{OS}/2V_T}$$

where r_e is the emitter resistance of each of Q_1 to Q_6, and R is the nominal value of R_1 and R_2. (*Hint:* Use Eq. 13.94.)

(b) Find $\Delta R/R$ to trim a 3-mV offset to zero.

(c) What is the maximum offset voltage that can be trimmed this way (corresponding to R_2 completely shorted)? (Recall that each of Q_5 and Q_6 is biased at 9.5 μA.)

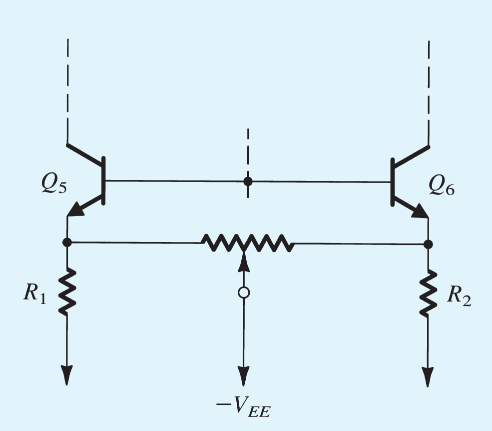


Figure P13.54

13.55 Through a processing imperfection, the β of Q_4 in Fig. 13.14 is reduced to 10, while the β of Q_3 remains at its regular value of 50. Assuming that the collector current of Q_3 remains unchanged at 9.5 μA, find the net output dc current of the Q_5–Q_6 mirror, and hence find also the input offset voltage that this mismatch introduces.

13.56 If the current transfer ratio of the mirror load of the 741 input stage is 0.995, find the CMRR of the input stage. (*Hint:* Use Eq. 13.102 together with the output resistance values determined in Exercise 13.28. Recall that the input-stage transistors are biased at 9.5 μA.)

13.57 Consider the circuit of Fig. 13.14 modified to include resistors R in series with the emitters of each of Q_8 and Q_9. What does the resistance looking into the collector of Q_9, R_{o9}, become? For what value of R does it equal R_{o10} (i.e., 31.1 MΩ)? For this case, what does R_o looking to the left of node Y become? (Recall that Q_9 is biased at 0.73 mA.)

***13.58** What is the effect on the differential gain of the 741 op amp of short-circuiting one or the other or both of R_1 and R_2 in Fig. 13.14? (Refer to Fig. 13.19.) For simplicity, assume $\beta = \infty$.

13.59 An alternative approach to that presented in Example 13.5 for determining the CMRR of the 741 input stage is investigated in this problem. Rather than performing the analysis on the closed loop shown in Fig. 13.23, we observe that the negative feedback increases the resistance at node Y by the amount of negative feedback. Thus, we can break the loop at Y and connect a resistance $R_f = (1+A\beta)R_o$ between the common-base connection of Q_3–Q_4 and ground. We can then determine the current i and G_{mcm}. Using the fact that the loop gain is approximately equal to β_P (Exercise 13.17) show that this approach yields an identical result to that found in Example 13.5.

13.60 Consider a variation on the design of the 741 second stage in which $R_8 = 50\ \Omega$. What R_{i2} and G_{m2} correspond?

D 13.61 In the analysis of the 741 second stage, note that R_{o2} is affected most strongly by the low value of R_{o13B}. Consider the effect of placing appropriate resistors in the emitters of Q_{12}, Q_{13A}, and Q_{13B} on this value. What resistor in the emitter of Q_{13B} would be required to make R_{o13B} equal to R_{o17} and thus R_{o2} half as great? What resistors in each of the other emitters would be required?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

13.62 For a 741 employing ±5-V supplies, $|V_{BE}| = 0.6$ V and $|V_{CEsat}| = 0.2$ V, find the output voltage limits that apply.

D 13.63 Consider an alternative to the present 741 output stage in which Q_{23} is not used, that is, in which its base and emitter are joined. Reevaluate the reflection of $R_L = 2$ kΩ to the collector of Q_{17}. What does A_2 become?

13.64 Figure P13.64 shows the circuit for determining the op-amp output resistance when v_O is positive and Q_{14} is conducting most of the current. Using the resistance of the Q_{18}–Q_{19} network calculated in Exercise 13.35 (163 Ω) and neglecting the large output resistance of Q_{13A}, find R_{out} when Q_{14} is sourcing an output current of 5 mA.

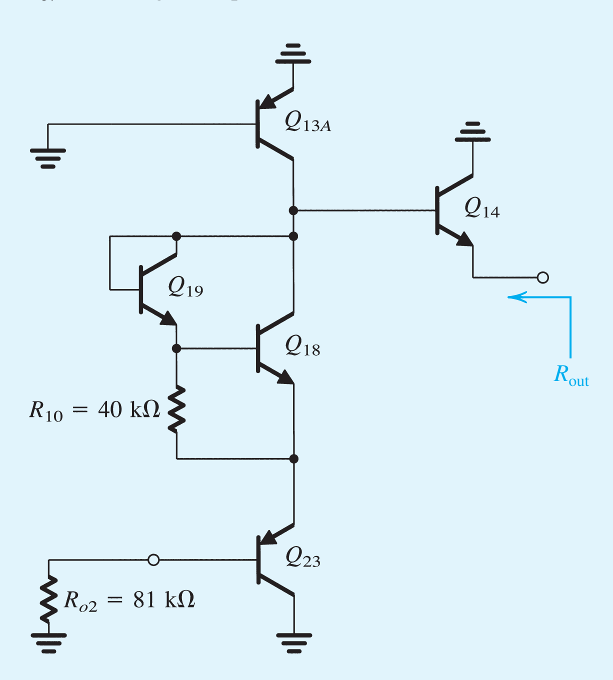


Figure P13.64

13.65 Consider the positive current-limiting circuit involving Q_{13A}, Q_{15}, and R_6. Find the current in R_6 at which the collector current of Q_{15} equals the current available from Q_{13A} (180 μA) minus the base current of Q_{14}. (You need to perform a couple of iterations.)

D 13.66 Consider the 741 sinking-current limit involving R_7, Q_{21}, Q_{24}, R_{11}, and Q_{22}. For what current through R_7 is the current in Q_{22} equal to the maximum current available from the input stage (i.e., the current in Q_8)? What simple change would you make to reduce this current limit to 10 mA?

13.67 Using the data provided in Eq. (13.115) (alone) for the overall gain of the 741 with a 2-kΩ load, and realizing the significance of the factor 0.97 in relation to the load, calculate the open-circuit voltage gain, the output resistance, and the gain with a load of 500 Ω.

13.68 A 741 op amp has a phase margin of 80°. If the excess phase shift is due to a second single pole, what is the frequency of this pole?

13.69 A 741 op amp has a phase margin of 80°. If the op amp has nearly coincident second and third poles, what is their frequency?

D *13.70 For a modified 741 whose second pole is at 5 MHz, what dominant-pole frequency is required for 85° phase margin with a closed-loop gain of 100? Assuming C_C continues to control the dominant pole, what value of C_C would be required?

13.71 An internally compensated op amp having an f_t of 5 MHz and dc gain of 10^6 utilizes Miller compensation around an inverting amplifier stage with a gain of –1000. If space exists for at most a 50-pF capacitor, what resistance level must be reached at the input of the Miller amplifier for compensation to be possible?

13.72 Consider the integrator op-amp model shown in Fig. 13.30. For $G_{m1} = 2$ mA/V, $C_C = 100$ pF, and a resistance of 2×10^7 Ω shunting C_C, sketch and label a Bode plot for the magnitude of the open-loop gain. If G_{m1} is related to the first-stage bias current as $G_{m1} = I/2V_T$, find the slew rate of this op amp.

13.73 For an amplifier with a slew rate of 10 V/μs, what is the full-power bandwidth for outputs of ±10 V? What unity-gain bandwidth, ω_t, would you expect if the topology was similar to that of the 741?

D 13.74 If a resistance R_E is included in each of the emitter leads of Q_3 and Q_4 of the 741 circuit, show that the slew rate is $4(V_T + IR_E/2)\omega_t$. Hence find the value of R_E that would double the 741 slew rate while keeping ω_t and I unchanged. What are the new values of C_C, the dc gain, and the 3-dB frequency?

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

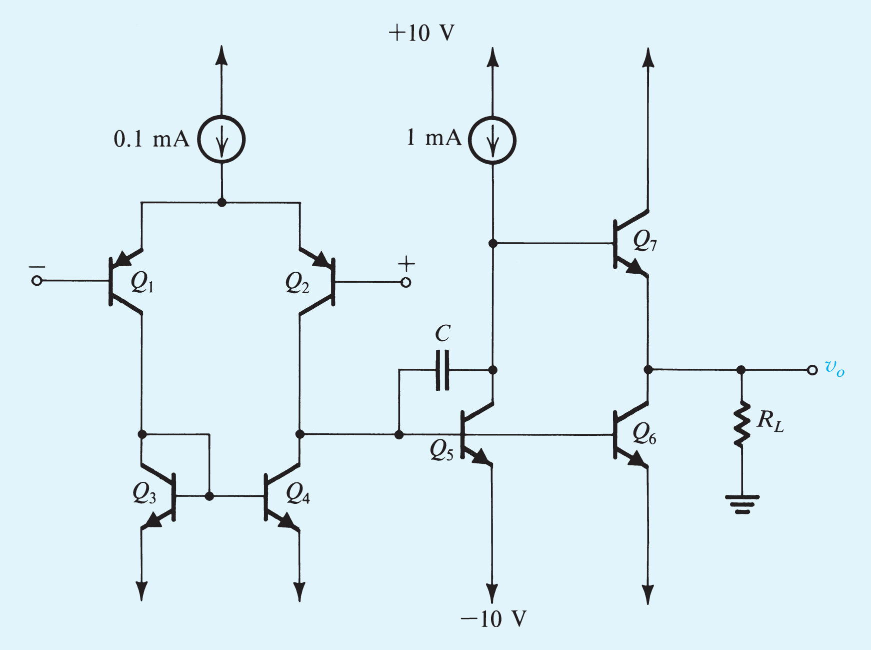


Figure P13.75

D 13.75 Figure P13.75 shows a circuit suitable for op-amp applications. For all transistors $\beta = 100$, $V_{BE} = 0.7$ V, and $r_o = \infty$.

(a) For inputs grounded and output held at 0 V (by negative feedback) find the collector currents of all transistors. Neglect base currents.
(b) Calculate the input resistance.
(c) Calculate the gain of the amplifier with a load of 5 kΩ.
(d) With load as in (c) calculate the value of the capacitor C required for a 3-dB frequency of 100 Hz.

Section 13.4: Modern Techniques for the Design of BJT Op Amps

Unless otherwise specified, for the problems in this section assume $\beta_N = 40$, $\beta_P = 10$, $V_{An} = 30$ V, $|V_{Ap}| = 20$ V, $|V_{BE}| = 0.7$ V, $|V_{CEsat}| = 0.1$ V.

D 13.76 Design the circuit in Fig. 13.33 to generate a current $I = 5\ \mu$A. Utilize transistors Q_1 and Q_2 having areas in a ratio of 1:4. Assume that Q_3 and Q_4 are matched and design for a 0.15-V drop across each of R_3 and R_4. Specify the values of R_2, R_3, and R_4. Ignore base currents.

D 13.77 Consider the circuit of Fig. 13.33 for the case designed in Exercise 13.37, namely, $I = 10\ \mu$A, $I_{S2}/I_{S1} = 2$, $R_2 = 1.73$ kΩ, $R_3 = R_4 = 20$ kΩ. Augment the circuit with *npn* transistors Q_5 and Q_6 with emitters connected to ground and bases connected to V_{BIAS1}, to generate constant currents of 10 µA and 40 µA, respectively. What should the emitter areas of Q_5 and Q_6 be relative to that of Q_1? What value of a resistance R_6 will, when connected in the emitter of Q_6, reduce the current generated by Q_6 to 10 µA? Assuming that the V_{BIAS1} line has a low incremental resistance to ground, find the output resistance of current source Q_5 and of current source Q_6 with R_6 connected. Ignore base currents.

D 13.78 It is required to use the circuit in Fig. 13.33 to bias an *npn* differential pair. The bias current-source transistor of the pair, Q_5, is identical to Q_2 and its base is connected to the BIAS1 line. In its emitter lead is connected a resistance R_5 equal to R_2. The differential pair has two equal collector

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

resistances R_C connected to V_{CC}, and the output voltage v_o is taken between the two collectors.

(a) Find an expression for the differential gain A_d in terms of (R_C/R_5) and (I_{S5}/I_{S1}). Comment on the expected temperature dependence of A_d. Neglect the effect of finite β_N.

(b) Design the circuit for $I = 20\ \mu A$ and $A_d = 10$ V/V. Let the emitter areas of Q_1 and Q_5 be in the ratio 1:4. Specify the required values of R_5 and R_C.

D 13.79 (a) Find the input common-mode range of the circuit in Fig. 13.35(a). Let $V_{CC} = 3$ V and $V_{\text{BIAS}} = 2.3$ V.
(b) Give the complementary version of the circuit in Fig. 13.35(a), that is, the one in which the differential pair is *npn*. For the same conditions as in (a), what is the input common-mode range?

13.80 For the circuit in Fig. 13.35(b), let $V_{CC} = 3$ V, $V_{\text{BIAS}} = 2.3$ V, $I = 20\ \mu A$, and $R_C = 25$ kΩ. Find the input common-mode range and the differential voltage gain v_o/v_{id}. Neglect base currents.

D 13.81 For the circuit in Fig. 13.36, let $V_{CC} = 3$ V, $V_{\text{BIAS}} = 0.7$ V, and $I_{C6} = 40\ \mu A$. Find R_C that results in a differential gain of 10 V/V. What is the input common-mode range and the input differential resistance? Ignore base currents except when calculating R_{id}. If R_{id} is to be increased by a factor of 4 while the gain and V_{ICM} remain unchanged, what must I and R_C be changed to?

13.82 It is required to find the input resistance and the voltage gain of the input stage shown in Fig. 13.37. Let $V_{ICM} \ll 0.8$ V so that the Q_3–Q_4 pair is off. Assume that Q_5 supplies 8 μA, that each of Q_7 to Q_{10} is biased at 8 μA, and that all four cascode transistors are operating in the active mode. The input resistance of the second stage of the op amp is 1.5 MΩ. The emitter-degeneration resistances are $R_7 = R_8 = 22$ kΩ, and $R_9 = R_{10} = 33$ kΩ. [*Hint*: Refer to Fig. 13.38.]

D *13.83 Consider the equivalent half-circuit shown in Fig. 13.38. Assume that in the original circuit, Q_1 is biased at a current I, Q_7 and Q_9 are biased at $2I$, the dc voltage drop across R_7 is 0.2 V, and the dc voltage drop across R_9 is 0.3 V. Find the output resistance in terms of I, and hence find the open-circuit voltage gain (i.e., the voltage gain for $R_L = \infty$). Now with R_L connected, find the voltage gain in terms of (IR_L). For $R_L = 1\,\text{M}\Omega$, find I that will result in the voltage gains of 150 V/V and 300 V/V.

***13.84** (a) For the circuit in Fig. 13.39, show that the loop gain of the common-mode feedback loop is

$$A\beta \simeq \frac{R_{o9} \parallel R_{o7}}{r_{e7} + R_7}$$

Recall that the CMF circuit responds only to the average voltage V_{CM} of its two input voltages and realizes the transfer characteristic $V_B = V_{CM} + 0.4$. Ignore the loading effect of the CMF circuit on the collectors of the cascode transistors.
(b) For the values in Example 13.8, calculate the loop gain $A\beta$.
(c) In Example 13.8, we found that with the CMF absent, a current mismatch $\Delta I = 0.3\ \mu A$ gives rise to $\Delta V_{CM} = 2.5$ V. Now, with the CMF present, use the value of loop gain found in (b) to calculate the expected ΔV_{CM} and compare to the value found by a different approach in Example 13.8. [*Hint*: Recall that negative feedback reduces change by a factor equal to $(1 + A\beta)$.]

13.85 The output stage in Fig. 13.41 operates at a quiescent current I_Q of 0.6 mA. The maximum current i_L that the stage can provide in either direction is 12 mA. Also, the output stage is equipped with a feedback circuit that maintains a minimum current of $I_Q/2$ in the inactive output transistor. Also, $V_{CC} = 3$ V.

(a) What is the allowable range of v_O?
(b) For $i_L = 0$, what is the output resistance of the op amp?
(c) If the open-loop gain of the op amp is 100,000 V/V, find the closed-loop output resistance obtained when the op amp is connected in the unity-gain voltage follower configuration, with $i_L = 0$.
(d) If the op amp is sourcing a load current $i_L = 12$ mA, find i_P, i_N, and the open-loop output resistance.
(e) Repeat (d) for the case of the open-loop op amp sinking a load current of 12 mA.

13.86 It is required to derive the expressions in Eqs. (13.130) and (13.131). Toward that end, first find v_{B7} in terms of v_{BEN} and hence i_N. Then find v_{B6} in terms of i_P. For the latter purpose note that Q_4 measures v_{EBP} and develops a current $i_4 = (v_{EBP} - v_{EB4})/R_4$. This current is supplied to the series connection of Q_5 and R_5 where $R_5 = R_4$. In the expression you

obtain for v_{B6}, use the relationship

$$\frac{I_{SP}}{I_{S4}} = \frac{I_{SN}}{I_{S5}}$$

to express v_{B6} in terms of i_P and I_{SN}. Now with v_{B6} and v_{B7} determined, find i_{C6} and i_{C7}.

13.87 It is required to derive the expression for v_E in Eq. (13.132). Toward that end, note from the circuit in Fig. 13.43 that $v_E = v_{EB7} + v_{BEN}$ and note that Q_N conducts a current i_N and Q_7 conducts a current i_{C7} given by Eq. (13.131).

D 13.88 For the output stage in Fig. 13.43, find the current I_{REF} that results in a quiescent current $I_Q = 0.6$ mA. Assume that $I = 12\ \mu$A, Q_N has eight times the area of Q_{10}, and Q_7 has four times the area of Q_{11}. What is the minimum current in Q_N and Q_P?

PART III

Digital Integrated Circuits

There are two indisputable facts about digital systems. They have dramatically changed our lives; and the digital revolution is driven by microelectronics.

Evidence of the pervasiveness and influence of digital systems can be found by thinking of what we do in our daily lives. Digital circuits exist in almost every electrical appliance we use in our homes; in the vehicles and transportation systems we use to travel; in the telephones and, most obviously, the mobile phones we use to communicate; in the medical equipment needed to care for our health; in the computers we use to do our work; and in the audio and video systems and the radio and TV sets we use to entertain ourselves. Indeed, it is very difficult to conceive of modern life without digital systems, none of which would have been possible without microelectronics.

Although the idea of a digital computing machine was conceived as early as the 1830s, early implementations were very cumbersome and used expensive mechanical devices. The first serious digital computers using vacuum tubes appeared in the 1930s and 1940s. These early computers used thousands of tubes and were housed literally in many rooms. Their fundamental limitation was low reliability: vacuum tubes had a finite life and needed large amounts of power. Had it not been for the invention of the transistor in 1947, ushering in the era of solid-state electronics, digital computers would have remained specialized machines used primarily in military and scientific applications.

By the mid-1950s, the first digital logic gates made of discrete bipolar transistors were commercially available. The invention of the integrated circuit in the late 1950s was also key, leading to the first digital IC in the early 1960s. Early digital ICs were made of bipolar transistors, with the most successful logic-circuit family of this type being transistor–transistor logic (or TTL), which dominated digital-circuit design, until the early 1980s.

Bipolar was replaced by NMOS, and NMOS by CMOS, again predominantly because of power dissipation and the need to pack more and more transistors on each IC chip. Bearing out Moore's law, which predicted in 1968 that IC chips would double the number of their transistors every two to three years (see Section 15.1), digital ICs have grown from a few transistors to 4 billion devices and to memory chips with 4-Gbit capacity.

Part III aims to provide a concise but nonetheless comprehensive and sufficiently detailed exposure to digital IC design. Our treatment is almost self-contained, requiring for the most part only a thorough understanding of the MOSFET material presented in Chapter 5. Thus Part III can be studied right after Chapter 5. The only exception to this is the last section in Chapter 15, which requires knowledge of the BJT (Chapter 6). Also, knowledge of the MOSFET internal capacitances (Section 10.2.2) will be needed.

Chapter 14 is the cornerstone of Part III. It provides a study of the bread-and-butter topic of digital IC design: the CMOS inverter and logic gates. Today, CMOS represents 98% of newly designed digital systems. The material in Chapter 14 is the minimum needed to learn something meaningful about digital circuits; it is a must study!

In Chapter 15, which has a modular structure, each section deals with an important though somewhat advanced topic. We consider the implications of technology scaling (Moore's law), survey digital IC technologies and logic-circuit families, and study the methodologies that make it possible to design complex digital IC chips with billions of transistors. We also present three types of MOS logic circuits that are used to supplement regular CMOS and look briefly at logic circuits that employ bipolar transistors.

Digital circuits can be broadly divided into logic and memory circuits. The latter form the subject of Chapter 16.

CHAPTER 14

CMOS Digital Logic Circuits

IN THIS CHAPTER YOU WILL LEARN

1. How CMOS logic circuits can be implemented by using arrangements of MOS transistors operating as voltage-controlled switches.
2. How to synthesize CMOS logic circuits that realize standard (e.g., inverter, NAND, and NOR gates) and complex Boolean functions.
3. How the operation of the logic inverter is characterized by such parameters as noise margins, propagation delay, and power dissipation; and how the inverter can be implemented by using one of three possible arrangements of voltage-controlled switches (transistors).
4. The structure, circuit operation, and design of the CMOS inverter, as well as static and dynamic performance analysis of the circuit.
5. How to select sizes for the transistors in a CMOS logic circuit so as to meet various performance requirements.
6. The sources of power consumption in logic circuits, with emphasis on CMOS, and the trade-off between power dissipation and speed of operation.

Introduction

This chapter provides a self-contained study of CMOS logic circuits, the bread and butter of digital IC design. We begin (Section 14.1) by learning how to synthesize CMOS circuits that implement various logic functions. This discussion will be at a high level without getting into the details of circuit operation and performance. To delve into these issues, we consider in Section 14.2 the most fundamental element of digital circuits: the logic inverter. We study its characterization, performance metrics, and methods of implementation. Out of this general study, CMOS emerges as the most ideal inverter implementation.

A thorough study of the CMOS inverter is undertaken in Section 14.3. This is followed by the consideration of the two most significant aspects of digital circuits: their speed of operation (Section 14.4) and their power consumption (Section 14.6). The fundamental design issue of selecting optimum sizes for the MOS transistors is dealt with in Section 14.5.

Besides presenting the most important digital IC technology (CMOS), this chapter lays the foundation for the more advanced topics studied in the two subsequent chapters.

14.1 CMOS Logic-Gate Circuits

In this section we consider the synthesis of CMOS circuits that realize combinational-logic functions. In combinational circuits, the output at any time is a function only of the values of input signals at that time. Thus, these circuits do not have memory and do not employ feedback. Combinational circuits are used in large quantities in every digital system.

14.1.1 Switch-Level Transistor Model

CMOS digital circuits utilize NMOS and PMOS transistors operating as switches. From Chapter 5, we know that a MOS transistor can operate as an on/off switch by using the gate voltage to operate the transistor in the triode region ("on" position) and in the cutoff region ("off" position).

Specifically, an NMOS transistor behaves as a closed switch, exhibiting a very small resistance (R_{on} or r_{DS}) between its drain and source terminals when its gate voltage is "high," usually at the power-supply level V_{DD}, which represents a logic 1. Conversely, when the gate voltage is "low" (i.e., at or close to ground voltage), which represents a logic 0, the transistor is cut off, thus conducting zero current and acting as an open switch. This is illustrated in Fig. 14.1(a).

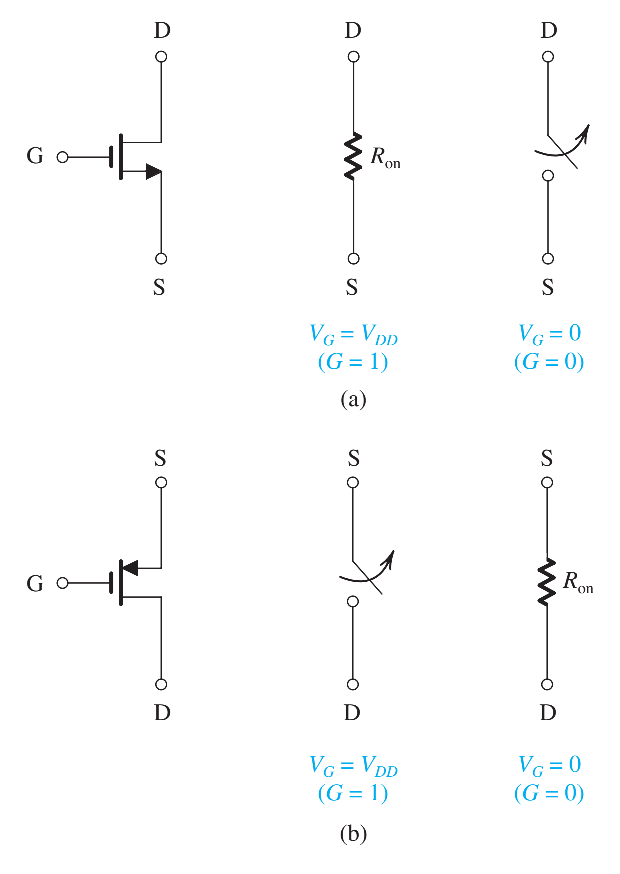


Figure 14.1 Operation of the (**a**) NMOS and (**b**) PMOS transistor as an on/off switch. The gate voltage controls the operation of the transistor switch, with the voltage V_{DD} representing a logic 1 and 0 V representing a logic 0. Note that the connections of the drain and source terminals are not shown.

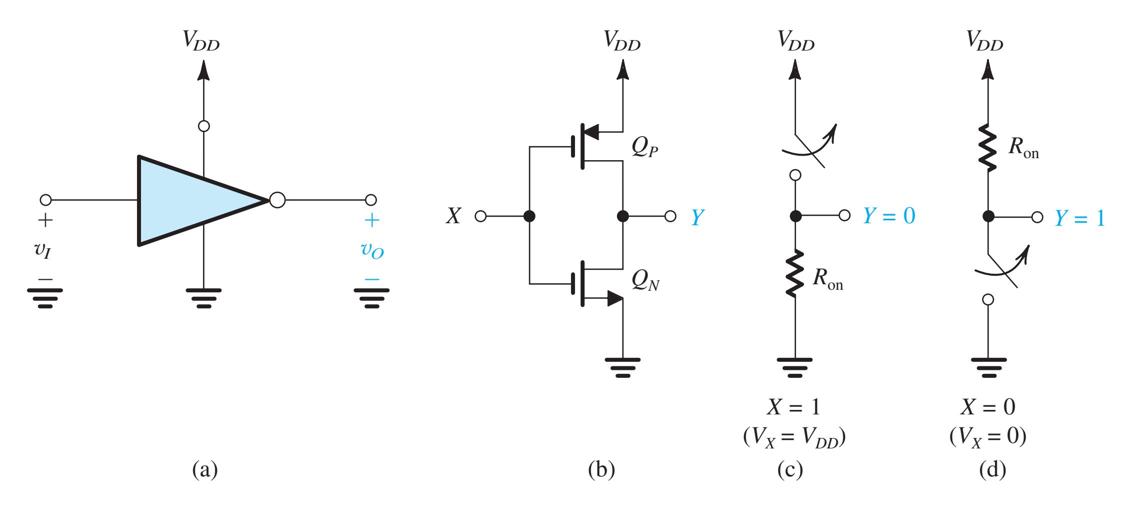


Figure 14.2 (**a**) Block representation of the logic inverter; (**b**) its CMOS realization; (**c**) operation when the input is a logic 1; (**d**) operation when the input is a logic 0.

The PMOS transistor operates in a complementary fashion: To turn the transistor on, its gate voltage is made low (0 V or logic 0). Raising the gate voltage to V_{DD} (logic 1) turns the PMOS transistor off. This is illustrated in Fig. 14.1(b).

We observe that the gate terminal of the MOSFET is used as the controlling node, and thus it is usually the input terminal of the logic gate.

14.1.2 The CMOS Inverter

Armed with this knowledge of the switching behavior of MOSFETs, let's consider making an inverter. As its name implies, the logic inverter inverts the logic value of its input signal. Thus, for a logic-0 input, the output will be a logic 1, and vice versa. Thus the logic function of the inverter can be represented by the Boolean expression

$$Y = \overline{X}$$

An inverter operated from a power supply V_{DD} is shown in block form in Fig. 14.2(a). Its CMOS circuit implementation is shown in Fig. 14.2(b). It consists of an NMOS transistor Q_N and a PMOS transistor Q_P, with the gate terminals connected together to constitute the inverter input terminal, to which a logic input X is applied. Also, both drain terminals are connected together to constitute the inverter output terminal on which the output logic variable Y appears.

When $X = 1$—that is, $V_X = V_{DD}$ [Fig. 14.2(c)]—the PMOS transistor will be off but the NMOS transistor will be on and will be connecting the inverter output terminal to ground through the small on-resistance R_{on}. Thus, the output voltage will be zero and $Y = 0$. When $X = 0$, that is, $V_X = 0$ [see Fig. 14.2(d)], the NMOS transistor will be off but the PMOS transistor will be on and will be connecting the output terminal to V_{DD} through the small resistance R_{on}. Thus the output voltage will be equal to V_{DD} and Y will be 1.

14.1.3 General Structure of CMOS Logic

A CMOS logic circuit is in effect an extension, or a generalization, of the CMOS inverter: The inverter consists of an NMOS **pull-down transistor** and a PMOS **pull-up transistor**,

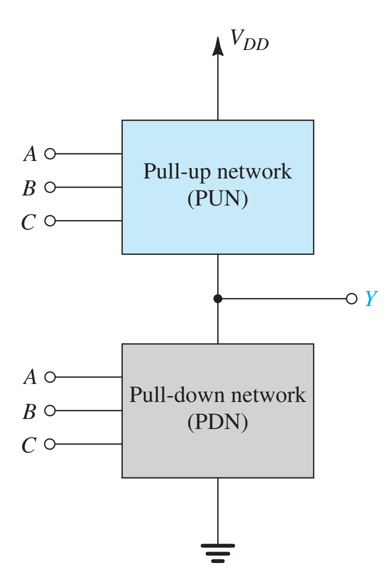


Figure 14.3 Representation of a three-input CMOS logic gate. The PUN comprises PMOS transistors, and the PDN comprises NMOS transistors.

operated by the input voltage in a complementary fashion. The CMOS logic gate consists of two networks: the **pull-down network (PDN)** constructed of NMOS transistors, and the **pull-up network (PUN)** constructed of PMOS transistors (see Fig. 14.3). The two networks are operated by the input variables, in a complementary fashion. Thus, for the three-input gate represented in Fig. 14.3, the PDN will conduct for all input combinations that require a low output ($Y = 0$) and will then pull the output node down to ground, causing a zero voltage to appear at the output, $v_Y = 0$. Simultaneously, the PUN will be off, and no direct dc path will exist between V_{DD} and ground. On the other hand, all input combinations that call for a high output ($Y = 1$) will cause the PUN to conduct, and the PUN will then pull the output node up to V_{DD}, establishing an output voltage $v_Y = V_{DD}$. Simultaneously, the PDN will be cut off, and again, no dc current path between V_{DD} and ground will exist in the circuit.

Now, since the PDN comprises NMOS transistors, and since an NMOS transistor conducts when the signal at its gate is high, the PDN is activated (i.e., conducts) when the inputs are high. In a dual manner, the PUN comprises PMOS transistors, and a PMOS transistor conducts when the input signal at its gate is low; thus the PUN is activated when the inputs are low.

The PDN utilizes devices in parallel to form an OR function and devices in series to form an AND function; the same is true of the PUN. Here, the OR and AND notation refer to current flow or conduction. Figure 14.4 shows examples of PDNs. For the circuit in Fig. 14.4(a), we observe that Q_A will conduct when A is high ($v_A = V_{DD}$) and will then pull the output node down to ground ($v_Y = 0$ V, $Y = 0$). Similarly, Q_B conducts and pulls Y down when B is high. Thus Y will be low when A is high *or* B is high, which can be expressed as

$$\overline{Y} = A + B$$

or equivalently

$$Y = \overline{A + B}$$

The PDN in Fig. 14.4(b) will conduct only when A and B are both high simultaneously. Thus Y will be low when A is high *and* B is high,

$$\overline{Y} = AB$$

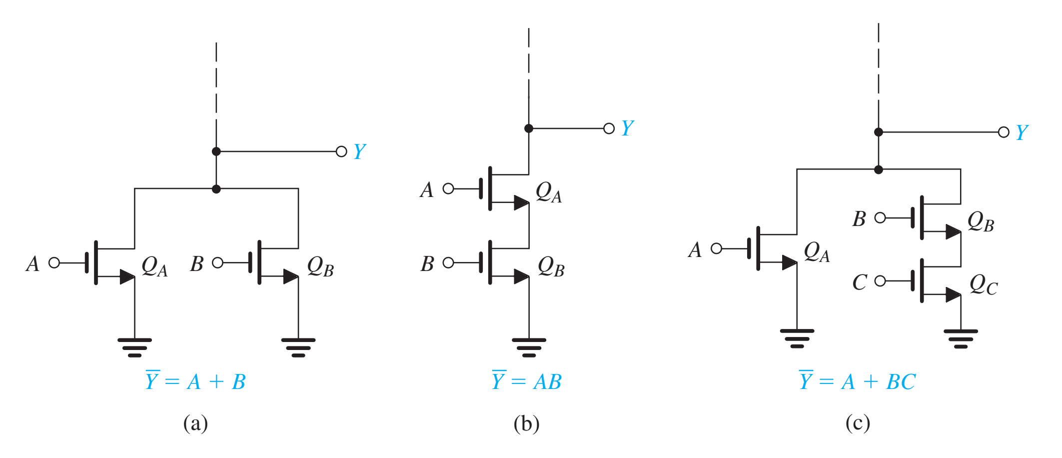


Figure 14.4 Examples of pull-down networks.

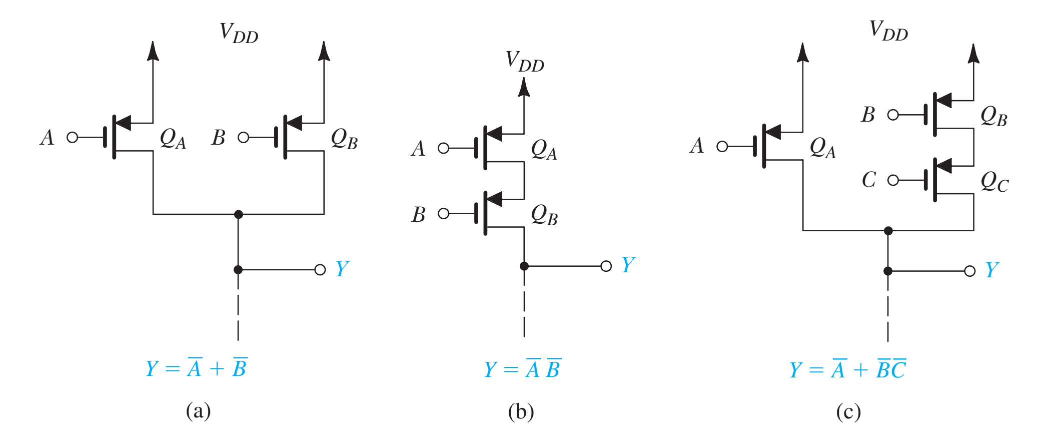


Figure 14.5 Examples of pull-up networks.

or equivalently

$$Y = \overline{AB}$$

As a final example, the PDN in Fig. 14.4(c) will conduct and cause Y to be 0 when A is high *or* when B *and* C are both high, thus

$$\overline{Y} = A + BC$$

or equivalently

$$Y = \overline{A + BC}$$

Next consider the PUN examples shown in Fig. 14.5. The PUN in Fig. 14.5(a) will conduct and pull Y up to $V_{DD}(Y = 1)$ when A is low *or when* B is low, thus

$$Y = \overline{A} + \overline{B}$$

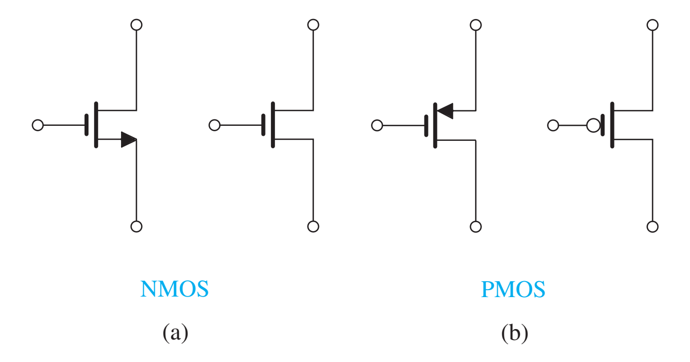


Figure 14.6 Usual and alternative circuit symbols for MOSFETs.

The PUN in Fig. 14.5(b) will conduct and produce a high output ($v_Y = V_{DD}$, $Y = 1$) only when *A and B* are both low, thus

$$Y = \overline{A}\,\overline{B}$$

Finally, the PUN in Fig. 14.5(c) will conduct and cause Y to be high (logic 1) if A is low *or* if *B and C* are both low; thus,

$$Y = \overline{A} + \overline{B}\,\overline{C}$$

Having developed an understanding and an appreciation of the structure and operation of PDNs and PUNs, we are almost ready to consider complete CMOS gates. First, however, we wish to introduce alternative circuit symbols that are almost universally used for MOS transistors by digital-circuit designers. Figure 14.6 shows our usual symbols (left) and the corresponding "digital" symbols (right). Observe that the symbol for the PMOS transistor with a circle at the gate terminal is intended to indicate that the signal at the gate has to be low for the device to be activated (i.e., to conduct). Thus, in terms of logic-circuit terminology, the gate terminal of the PMOS transistor is an *active low* input. Besides indicating this property of PMOS devices, the digital symbols omit any indication of which of the device terminals is the source and which is the drain. This should cause no difficulty at this stage of our study; simply remember that for an NMOS transistor, the drain is the terminal that is at the higher voltage (current flows from drain to source), and for a PMOS transistor the source is the terminal that is at the higher voltage (current flows from source to drain). To be consistent with the literature, we shall henceforth use these modified symbols for MOS transistors in logic applications, except in locations where our usual symbols help in understanding circuit operation.

14.1.4 The Two-Input NOR Gate

We first consider the CMOS gate that realizes the two-input NOR function

$$Y = \overline{A + B} = \overline{A}\,\overline{B} \tag{14.1}$$

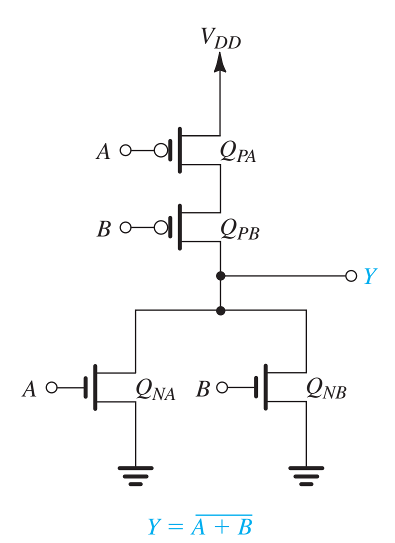


Figure 14.7 A two-input CMOS NOR gate.

We see that Y is to be low (PDN conducting) when A is high or B is high. Thus the PDN consists of two parallel NMOS devices with A and B as inputs [i.e., the circuit in Fig. 14.4(a)]. For the PUN, we note from the second expression in Eq. (14.1) that Y is to be high when A and B are both low. Thus the PUN consists of two series PMOS devices with A and B as the inputs [i.e., the circuit in Fig. 14.5(b)]. Putting the PDN and the PUN together gives the CMOS NOR gate shown in Fig. 14.7. Note that extension to a higher number of inputs is straightforward: For each additional input, an NMOS transistor is added in parallel with Q_{NA} and Q_{NB}, and a PMOS transistor is added in series with Q_{PA} and Q_{PB}.

14.1.5 The Two-Input NAND Gate

The two-input NAND function is described by the Boolean expression

$$Y = \overline{AB} = \overline{A} + \overline{B} \tag{14.2}$$

To synthesize the PDN, we consider the input combinations that require Y to be low: There is only one such combination, namely, A and B both high. Thus, the PDN simply comprises two NMOS transistors in series [such as the circuit in Fig. 14.4(b)]. To synthesize the PUN, we consider the input combinations that result in Y being high. These are found from the second expression in Eq. (14.2) as A low or B low. Thus, the PUN consists of two parallel PMOS transistors with A and B applied to their gates [such as the circuit in Fig. 14.5(a)]. Putting the PDN and PUN together results in the CMOS NAND gate implementation shown in Fig. 14.8. Note that extension to a higher number of inputs is straightforward: For each additional input, we add an NMOS transistor in series with Q_{NA} and Q_{NB}, and a PMOS transistor in parallel with Q_{PA} and Q_{PB}.

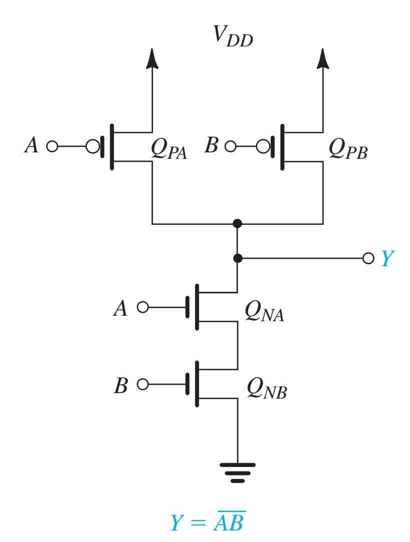


Figure 14.8 A two-input CMOS NAND gate.

14.1.6 A Complex Gate

Consider next the more complex logic function

$$Y = \overline{A(B + CD)} \tag{14.3}$$

Since $\overline{Y} = A(B + CD)$, we see that Y should be low for A high and simultaneously either B high or C and D both high, from which the PDN is directly obtained. To obtain the PUN, we need to express Y in terms of the complemented variables. We do this through repeated application of DeMorgan's law, as follows:

$$\begin{aligned} Y &= \overline{A(B + CD)} \\ &= \overline{A} + \overline{B + CD} \\ &= \overline{A} + \overline{B}\,\overline{CD} \\ &= \overline{A} + \overline{B}\left(\overline{C} + \overline{D}\right) \end{aligned} \tag{14.4}$$

Thus, Y is high for A low or B low and either C or D low. The corresponding complete CMOS circuit will be as shown in Fig. 14.9.

14.1.7 Obtaining the PUN from the PDN and Vice Versa

From the CMOS gate circuits considered thus far (e.g., that in Fig. 14.9), we observe that the PDN and the PUN are dual networks: Where a series branch exists in one, a parallel branch exists in the other. Thus, we can obtain one from the other, a process that can be simpler than having to synthesize each separately from the Boolean expression of the function. For instance, in the circuit of Fig. 14.9, we found it relatively easy to obtain the PDN, simply because we already had $\overline{Y}$ in terms of the uncomplemented inputs. On the other hand, to obtain the PUN, we had to manipulate the given Boolean expression to express Y as a function of the complemented variables, the form convenient for synthesizing PUNs. Alternatively, we could have used this duality property to obtain the PUN from the PDN. The reader is urged to refer to Fig. 14.9 to convince herself that this is indeed possible.

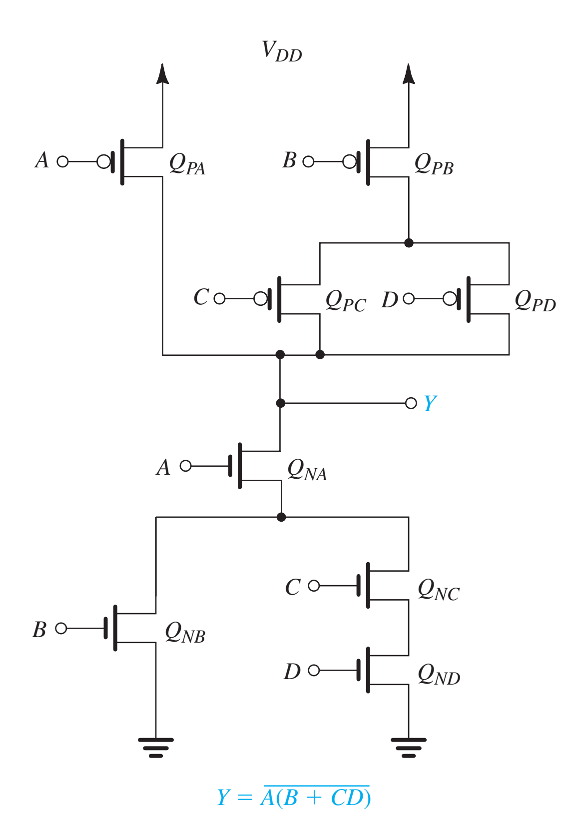


Figure 14.9 CMOS realization of a complex gate.

It should, however, be mentioned that at times it is not easy to obtain one of the two networks from the other using the duality property. For such cases, one has to resort to a more rigorous process, which is beyond the scope of this book (see Kang and Leblebici, 1999).

14.1.8 The Exclusive-OR Function

An important function that often arises in logic design is the exclusive-OR (XOR) function,

$$Y = A\overline{B} + \overline{A}B \tag{14.5}$$

We observe that since Y (rather than $\overline{Y}$) is given, it is easier to synthesize the PUN. We note, however, that unfortunately Y is not a function of the complemented variables only (as we would like it to be). Thus, we will need additional inverters. The PUN obtained directly from Eq. (14.5) is shown in Fig. 14.10(a). Note that the Q_1, Q_2 branch realizes the first term ($A\overline{B}$), whereas the Q_3, Q_4 branch realizes the second term ($\overline{A}B$). Note also the need for two additional inverters to generate $\overline{A}$ and $\overline{B}$.

As for synthesizing the PDN, we can obtain it as the dual network of the PUN in Fig. 14.10(a). Alternatively, we can develop an expression for $\overline{Y}$ and use it to synthesize the PDN. Leaving the first approach for the reader to do as an exercise, we shall utilize the direct synthesis approach. DeMorgan's law can be applied to the expression in Eq. (14.5) to obtain $\overline{Y}$ as

$$\overline{Y} = AB + \overline{A}\,\overline{B} \tag{14.6}$$

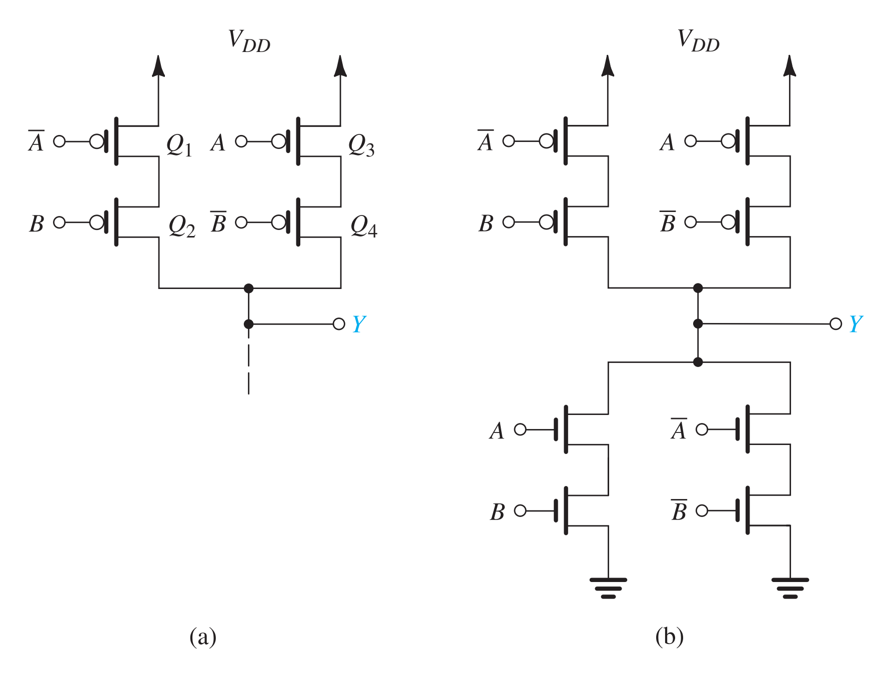


Figure 14.10 Realization of the exclusive-OR (XOR) function. **(a)** The PUN synthesized directly from the expression in Eq. (14.5). **(b)** The complete XOR realization utilizing the PUN in **(a)** and a PDN that is synthesized directly from the expression in Eq. (14.6). Note that two inverters (not shown) are needed to generate the complemented variables. Also note that in this XOR realization, the PDN and the PUN are not dual networks; however, a realization based on dual networks is possible (see Problem 14.9).

The corresponding PDN will be as in Fig. 14.10(b), which shows the CMOS realization of the exclusive-OR function except for the two additional inverters. Note that the exclusive-OR requires 12 transistors for its realization, a rather complex network. Later, in Section 15.5, we shall show a simpler realization of the XOR employing a different form of CMOS logic.

Another interesting observation follows from the circuit in Fig. 14.10(b). The PDN and the PUN here are *not* dual networks. Indeed, duality of the PDN and the PUN is not a necessary condition. Thus, although a dual of PDN (or PUN) can always be used for PUN (or PDN), the two networks are not necessarily duals.

14.1.9 Summary of the Synthesis Method

1. The PDN can be most directly synthesized by expressing $\overline{Y}$ as a function of the *uncomplemented* variables. If complemented variables appear in this expression, additional inverters will be required to generate them.
2. The PUN can be most directly synthesized by expressing Y as a function of the *complemented* variables and then applying the uncomplemented variables to the gates of the PMOS transistors. If uncomplemented variables appear in the expression, additional inverters will be needed.
3. The PDN can be obtained from the PUN (and vice versa) using the duality property.

Example 14.1

Synthesize a CMOS logic circuit that implements the Boolean function

$$Y = \overline{A + B(C + D)}$$

Solution

To obtain the PDN we use

$$\overline{Y} = A + B(C + D)$$

and note that Y will be low when A is high or when $[B(C + D)]$ is high. Thus we have two parallel networks. One consists of a single transistor with A supplied to its gate, and the second is a network composed of a transistor controlled by B in series with two parallel transistors controlled by C and D. The resulting PDN is shown in Fig. 14.11(a).

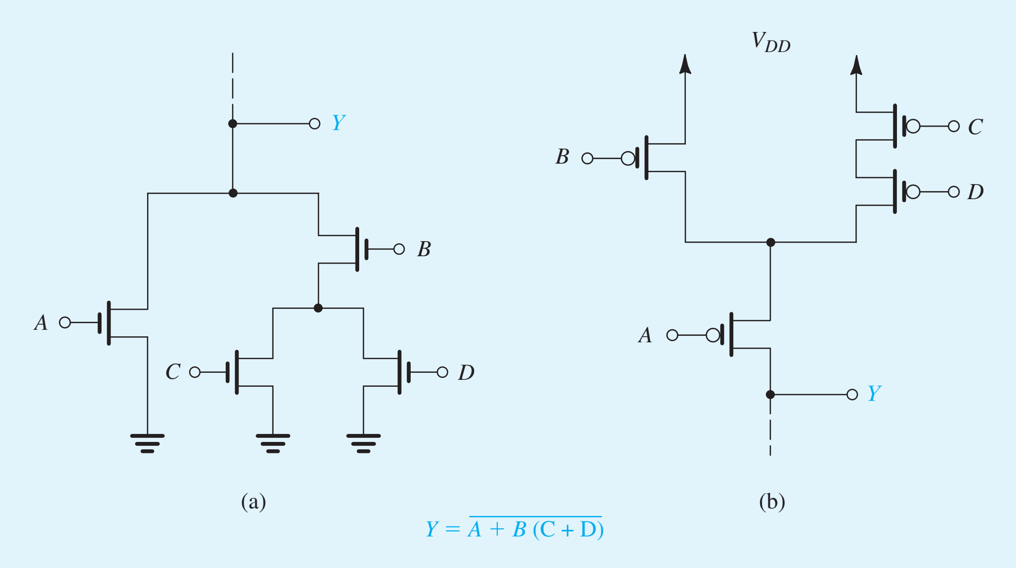


Figure 14.11 **(a)** The PDN and **(b)** the PUN for the logic function in Example 14.1.

A question arises: Should the B transistor be placed closer to ground (with the C and D transistors closer to Y) or closer to Y? From a logic standpoint, *both* are valid solutions. When one is optimizing for time delay, however, there may be a difference. We will explore this point in greater detail later on.

To obtain the PUN we use deMorgan's law on the given expression for Y to obtain

$$Y = \overline{A}(\overline{B} + \overline{C}\,\overline{D})$$

which leads to the circuit in Fig. 14.11(b). Here again, we may exchange the positions of the C and D transistors with no effect on the logic behavior. Finally, the complete circuit can be obtained by combining the PDN and the PUN shown in Fig. 14.11.

14.2 Digital Logic Inverters

Having learned how to synthesize CMOS circuits that implement various logic functions, we next consider the design and performance evaluation of these CMOS logic circuits. Toward that end, we step back to study in detail the characterization of the basic logic element, the inverter. We will do this in this section in general terms; that is, our study will not be confined to CMOS inverters only. In subsequent sections we specialize what we will have learned in this section to the case of the CMOS inverter and extend it to CMOS logic gates.

14.2.1 The Voltage-Transfer Characteristic (VTC)

Refer to the inverter shown in block form in Fig. 14.1(a). To quantify the operation of the inverter, we utilize its voltage-transfer characteristic (VTC). We have already introduced the concept of the VTC and utilized it to characterize the operation of basic MOSFET amplifiers in Section 7.1.3. Figure 14.12 shows such a circuit, together with its VTC. Observe that the circuit in fact implements the inverter function: For a logic-0 input, v_I is close to 0 V and specifically lower than the MOSFET threshold voltage V_{tn}, the transistor will be off, $i_D = 0$, and $v_O = V_{DD}$, which is a logic 1. For a logic-1 input, $v_I = V_{DD}$, the transistor will be conducting and operating in the triode region (at point D on the VTC), and the output voltage will be low (logic 0).

Thus to use this amplifier as a logic inverter, we utilize its extreme regions of operation. This is exactly the opposite to its use as a signal amplifier, where it would be biased at the middle of the transfer characteristic segment BC and the signal kept small enough to restrict operation to a short, almost linear, segment of the transfer curve. Digital applications, on the other hand, make use of the gross nonlinearity exhibited by the VTC.

With these observations in mind, we show in Fig. 14.13 a possible VTC of a logic inverter. For simplicity, we are using three straight lines to approximate the VTC, which is usually a

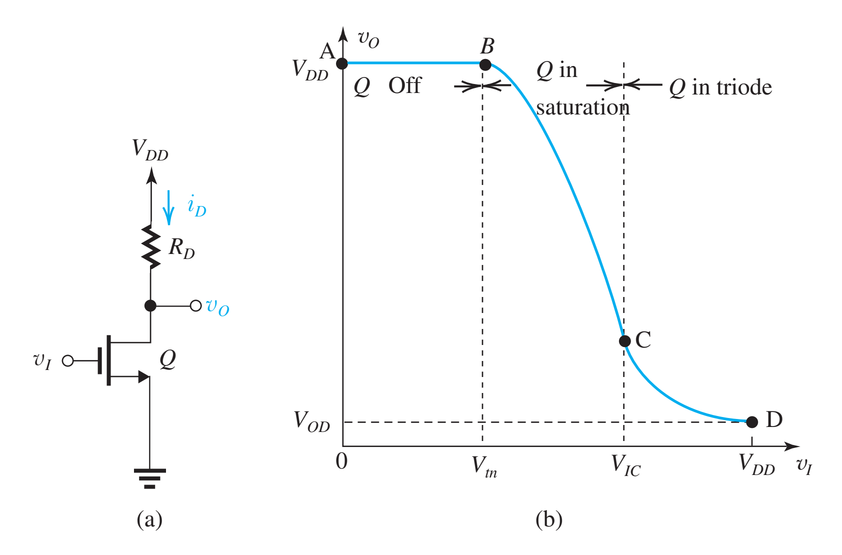


Figure 14.12 The simple resistively loaded MOS amplifier can be used as a logic inverter when operated in cutoff $(v_I < V_{tn})$ and in triode $(v_I > V_{IC})$. The output high level is V_{DD} and the low level is V_{OD}.

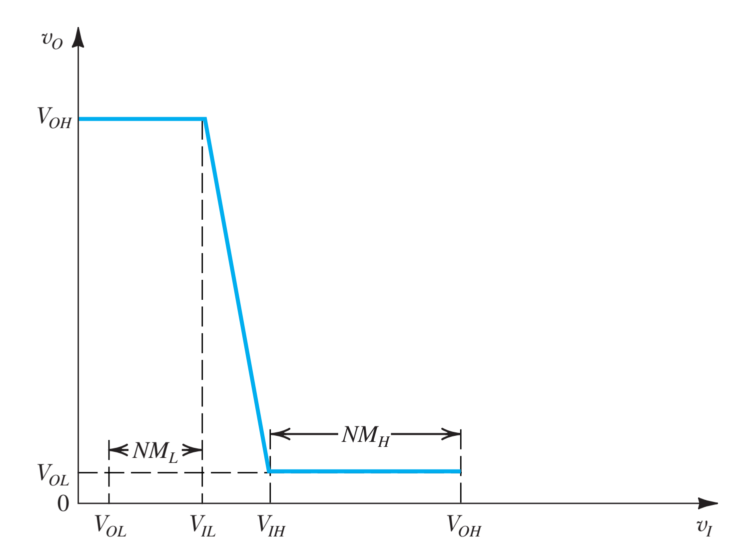


Figure 14.13 Voltage-transfer characteristic of an inverter. The VTC is approximated by three straight-line segments. Note the four parameters of the VTC (V_{OH}, V_{OL}, V_{IL}, and V_{IH}) and their use in determining the noise margins (NM_H and NM_L).

nonlinear curve such as that in Fig. 14.12. Observe that the output high level, denoted V_{OH}, does not depend on the exact value of v_I as long as v_I does not exceed the value labeled V_{IL}; when v_I exceeds V_{IL}, the output decreases and the inverter enters its amplifier region of operation, also called the **transition region**. It follows that V_{IL} is an important parameter of the inverter VTC: It is the *maximum value that v_I can have while being interpreted by the inverter as representing a logic 0.*

Similarly, we observe that the output low level, denoted V_{OL}, does not depend on the exact value of v_I as long as v_I does not fall below V_{IH}. Thus V_{IH} is an important parameter of the inverter VTC: It is the *minimum value that v_I can have while being interpreted by the inverter as representing a logic 1.*

14.2.2 Noise Margins

The insensitivity of the inverter output to the exact value of v_I within allowed regions is a great advantage that digital circuits have over analog circuits. To quantify this insensitivity property, consider the situation that occurs often in a digital system where an inverter (or a logic gate based on the inverter circuit) is driving another similar inverter, as shown in Fig. 14.14.

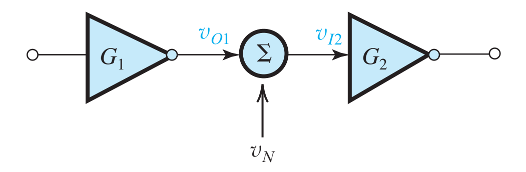


Figure 14.14 Noise voltage v_N is coupled to the interconnection between the output of inverter G_1 and the input of inverter G_2.

Table 14.1 Important Parameters of the VTC of the Logic Inverter (Refer to Fig. 14.13)

V_{OL}:	Output low level
V_{OH}:	Output high level
V_{IL}:	Maximum value of input interpreted by the inverter as a logic 0
V_{IH}:	Minimum value of input interpreted by the inverter as a logic 1
NM_L:	Noise margin for low input $= V_{IL} - V_{OL}$
NM_H:	Noise margin for high input $= V_{OH} - V_{IH}$

Here we assume that a noise or interference signal v_N is somehow coupled to the interconnection between the output of inverter G_1 and the input of inverter G_2 with the result that the input of G_2 becomes

$$v_{I2} = v_{O1} + v_N \tag{14.7}$$

where the noise voltage v_N can be either positive or negative. Now consider the case $v_{O1} = V_{OL}$; that is, inverter G_2 is driven by a logic-0 signal. Reference to Fig. 14.13 indicates that in this case G_2 will continue to function properly as long as its input v_{I2} does not exceed V_{IL}. Equation (14.7) then indicates that v_N can be as high as $V_{IL} - V_{OL}$ while G_2 continues to function properly. Thus, we can say that inverter G_2 has a **noise margin for low input**, NM_L, of

$$NM_L = V_{IL} - V_{OL} \tag{14.8}$$

Similarly, if $v_{O1} = V_{OH}$, the driven inverter G_2 will continue to see a high input as long as v_{I2} does not fall below V_{IH}. Thus, in the high-input state, inverter G_2 can tolerate a negative v_N of magnitude as high as $V_{OH} - V_{IH}$. We can thus state that G_2 has a **high-input noise margin**, NM_H, of

$$NM_H = V_{OH} - V_{IH} \tag{14.9}$$

In summary, four parameters, V_{OH}, V_{OL}, V_{IH}, and V_{IL}, define the VTC of an inverter and determine its noise margins, which in turn measure the ability of the inverter to tolerate variations in the input signal levels. In this regard, observe that changes in the input signal level within the noise margins are *rejected* by the inverter. Thus noise is not allowed to propagate further through the system, a definite advantage of digital over analog circuits. Alternatively, we can think of the inverter as *restoring* the signal levels to standard values (V_{OL} and V_{OH}) even when it is presented with corrupted input signal levels (within the noise margins). As a summary, useful for future reference, we present a listing and definitions of the important parameters of the inverter VTC in Table 14.1.

The formal definitions of the threshold voltages V_{IL} and V_{IH} are given in Fig. 14.15. Observe that V_{IL} and V_{IH} are defined as the VTC points at which the slope is -1 V/V. As v_I exceeds V_{IL}, the magnitude of the inverter gain increases and the VTC enters its transition region. Similarly, as v_I falls below V_{IH}, the inverter enters the transition region and the magnitude of the gain increases. Finally, note that Fig. 14.15 shows the definition of another important point on the VTC; this is point M at which $v_O = v_I$. Point M is loosely considered to be the midpoint of the VTC and thus the point at which the *inverter switches from one state to the other*. Point M plays an important role in the definition of the time delay of the inverter, as we shall see shortly.

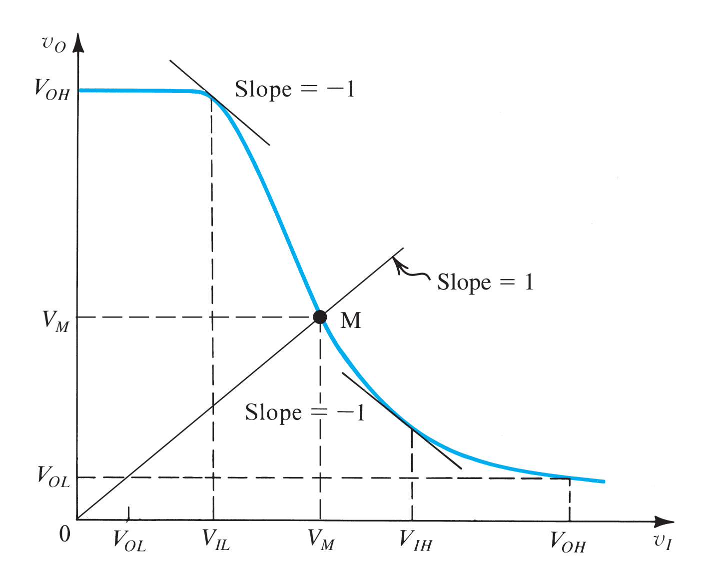


Figure 14.15 Typical voltage-transfer characteristic (VTC) of a logic inverter, illustrating the definition of the critical points.

14.2.3 The Ideal VTC

What constitutes an ideal VTC for an inverter? The answer to this naturally arising question follows directly from the preceding discussion: An ideal VTC is one that maximizes the output signal swing and the noise margins. For an inverter operated from a power supply V_{DD}, maximum signal swing is obtained when

$$V_{OH} = V_{DD}$$

and

$$V_{OL} = 0$$

To obtain maximum noise margins, we first arrange for the transition region to be made as narrow as possible and ideally of zero width. Then, the two noise margins are equalized by arranging for the transition from high to low to occur at the midpoint of the power supply, that is, at $V_{DD}/2$. The result is the VTC shown in Fig. 14.16, for which

$$V_{IL} = V_{IH} = V_M = V_{DD}/2$$

Observe that the sharp transition at $V_{DD}/2$ indicates that if the inverter were to be used as an amplifier, its gain would be infinite. Again, we point out that while the analog designer's interest would be focused on the transition region of the VTC, the digital designer would prefer the transition region to be as narrow as possible, as is the case in the ideal VTC of Fig. 14.16. Finally, we will see in Section 14.3 that inverters implemented using CMOS technology come very close to realizing the ideal VTC.

14.2.4 Inverter Implementation

Inverters are implemented using transistors (Chapters 5 and 6) operating as **voltage-controlled switches**. The simplest inverter implementation is shown in Fig. 14.17(a). The switch is

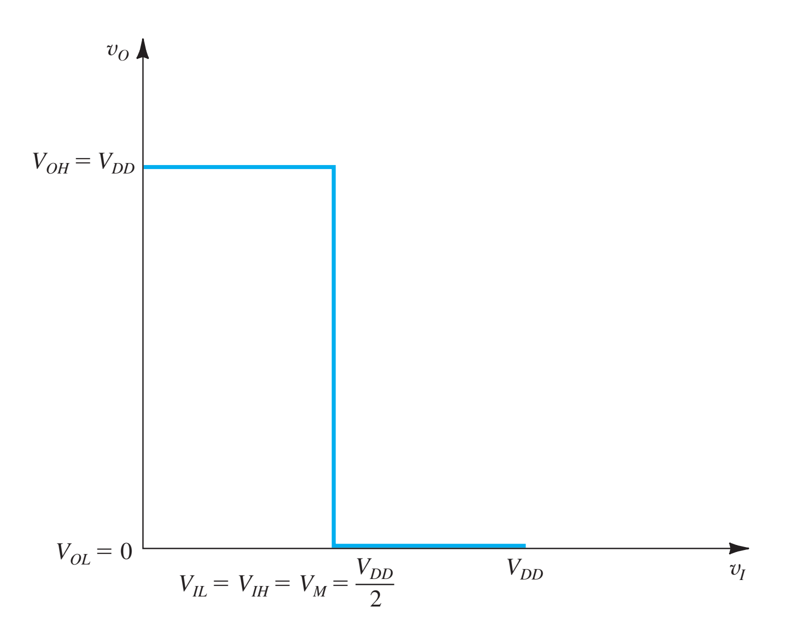


Figure 14.16 The VTC of an ideal inverter.

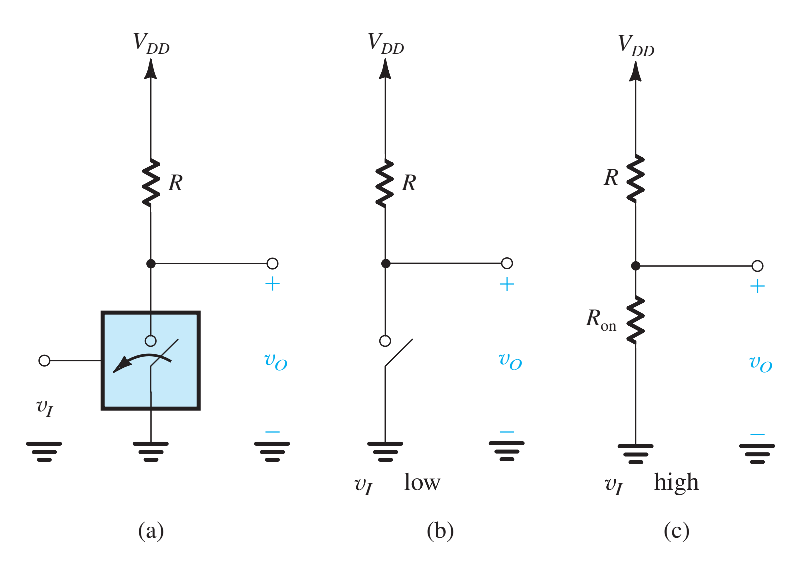


Figure 14.17 **(a)** The simplest implementation of a logic inverter using a voltage-controlled switch; **(b)** equivalent circuit when v_I is low; **(c)** equivalent circuit when v_I is high. Note that the switch is assumed to close when v_I is high.

controlled by the inverter input voltage v_I: When v_I is low, the switch will be open and $v_O = V_{DD}$, since no current flows through R. When v_I is high, the switch will be closed and, assuming an ideal switch, v_O will be 0.

Transistor switches, however, as we know from Chapters 5 and 6, are not perfect. Although their **off-resistances** are very high, and thus an open switch closely approximates an open circuit, the "on" switch has a finite closure, or **on-resistance**, R_{on}. The result is that when v_I

is high, the inverter has the equivalent circuit shown in Fig. 14.17(c), from which V_{OL} can be found.[1]

$$V_{OL} = V_{DD}\frac{R_{\text{on}}}{R + R_{\text{on}}}$$

We observe that the circuit in Fig. 14.12(a) is a direct implementation of the inverter in Fig. 14.17. In this case, R_{on} is equal to r_{DS} of the MOSFET evaluated at its operating point in the triode region with $V_{GS} = V_{DD}$.

EXERCISE

D14.1 Design the inverter in Fig. 14.12(a) to provide $V_{OL} = 90$ mV and to draw a supply current of 30 μA in the low-output state. Let the transistor be specified to have $V_t = 0.4$ V, $\mu_n C_{ox} = 125\ \mu\text{A/V}^2$, and $\lambda = 0$. The power supply $V_{DD} = 1.8$ V. Specify the required values of W/L and R_D. How much power is drawn from V_{DD} when the switch is open? Closed?
Hint: Recall that for small v_{DS},

$$r_{DS} \simeq 1\Big/\left[(\mu_n C_{ox})\left(\frac{W}{L}\right)(V_{GS} - V_t)\right]$$

Ans. 1.9; 57 kΩ; 0; 54 μW

More elaborate implementations of the logic inverter exist, and we show two of these in Fig. 14.18(a) and 14.19. The circuit in Fig. 14.18(a) utilizes a pair of **complementary switches**, the **"pull-up" (PU) switch** connects the output node to V_{DD}, and the **"pull-down" (PD) switch** connects the output node to ground. When v_I is low, the PU switch will be closed and the PD switch open, resulting in the equivalent circuit of Fig. 14.18(b). Observe that in this case R_{on} of PU connects the output to V_{DD}, thus establishing $V_{OH} = V_{DD}$. Also observe that no current flows, and thus no power is dissipated, in the circuit. Next, if v_I is raised to the logic-1 level, the PU switch will open while the PD switch will close, resulting in the equivalent circuit shown in Fig. 14.18(c). Here R_{on} of the PD switch connects the output to ground, thus establishing $V_{OL} = 0$. Here again no current flows, and no power is dissipated. The superiority of this inverter implementation over that using the single pull-down switch and a resistor (known as a **pull-up resistor**) should be obvious: With $V_{OL} = 0$ and $V_{OH} = V_{DD}$, the signal swing is at its maximum possible, and the power dissipation is zero in both states. This circuit constitutes the basis of the CMOS inverter that we synthesized in the previous section [Fig. 14.2(b)] and will study in detail in Section 14.3.

---

[1]If a BJT is used to implement the switch in Fig. 14.17(a), its equivalent circuit in the closed position includes in addition to the resistance $R_{\text{on}} = R_{CE\text{sat}}$ an offset voltage of about 50 mV to 100 mV [see Fig. 6.20(c)]. We shall not pursue this subject any further here, since the relatively long delay time needed to turn off a saturated BJT has caused the use of BJT switches operated in saturation to all but disappear from the digital IC world.

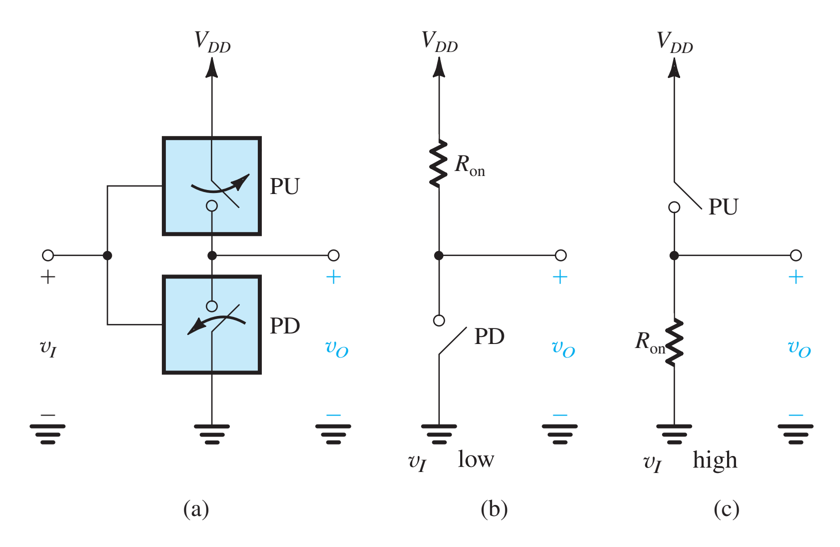


Figure 14.18 A more elaborate implementation of the logic inverter utilizing two complementary switches. This is the basis of the CMOS inverter that we synthesized in the previous section [Fig. 14.2(b)] and shall study in Section 14.3.

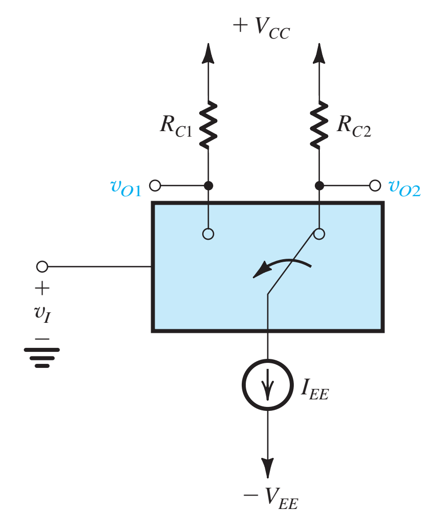


Figure 14.19 Another inverter implementation utilizing a double-throw switch to steer the constant current I_{EE} to R_{C1} (when v_I is high) or R_{C2} (when v_I is low). This is the basis of the emitter-coupled logic (ECL) studied briefly in Chapter 15.

Finally, consider the inverter implementation of Fig. 14.19. Here a double-throw switch is used to steer the constant current I_{EE} into one of two resistors connected to the positive supply V_{CC}. The reader is urged to show that if a high v_I results in the switch being connected to R_{C1}, then a logic inversion function is realized at v_{O1}. Note that the output voltage is independent of the switch resistance. This *current-steering* or *current-mode* logic arrangement is the basis of the fastest available digital logic circuits, called emitter-coupled logic (ECL), which we shall study briefly in Section 15.6.1. In fact, ECL is the only BJT logic-circuit type that is currently employed in new designs and the only one studied in this book.

EXERCISE

14.2 For the current-steering circuit in Fig. 14.19, let $V_{CC} = 5$ V, $I_{EE} = 1$ mA, and $R_{C1} = R_{C2} = 2$ kΩ. What are the high and low logic levels obtained at the outputs?
Ans. $V_{OH} = 5$ V; $V_{OL} = 3$ V

Example 14.2 Resistively Loaded MOS Inverter

For the simple MOS inverter in Fig. 14.12(a):

(a) Derive expressions for V_{OH}, V_{OL}, V_{IL}, V_{IH}, and V_M. For simplicity, neglect channel-length modulation (i.e., assume $\lambda = 0$). Show that these inverter parameters can be expressed in terms of V_{DD}, V_t, and $(k_n R_D)$. The latter parameter has the dimension of V^{-1}, and to simplify the expressions, denote $k_n R_D \equiv 1/V_x$.
(b) Show that V_x can be used as a design parameter for the inverter circuit. In particular, find the value of V_x that results in $V_M = V_{DD}/2$.
(c) Find numerical values for all parameters and for the inverter noise margins for $V_{DD} = 1.8$ V, $V_t = 0.5$ V, and V_x set to the value found in (b).
(d) For $k_n' = 300\ \mu A/V^2$ and $W/L = 1.5$, find the required value of R_D and use it to determine the average power dissipated in the inverter, assuming that the inverter spends half of the time in each of its two states.
(e) Comment on the characteristics of this inverter circuit vis-à-vis the ideal characteristics as well as on its suitability for implementation in integrated-circuit form.

Solution

(a) Refer to Fig. 14.20. For $v_I < V_t$, the MOSFET is off, $i_D = 0$, and $v_O = V_{DD}$. Thus

$$V_{OH} = V_{DD} \tag{14.10}$$

As v_I exceeds V_t, the MOSFET turns on and operates initially in the saturation region. Assuming $\lambda = 0$,

$$i_D = \frac{1}{2}k_n(v_I - V_t)^2$$

and

$$v_O = V_{DD} - R_D i_D = V_{DD} - \frac{1}{2}k_n R_D (v_I - V_t)^2$$

substituting $k_n R_D = 1/V_x$, the BC segment of the VTC is described by

$$v_O = V_{DD} - \frac{1}{2V_x}(v_I - V_t)^2 \tag{14.11}$$

Example 14.2 *continued*

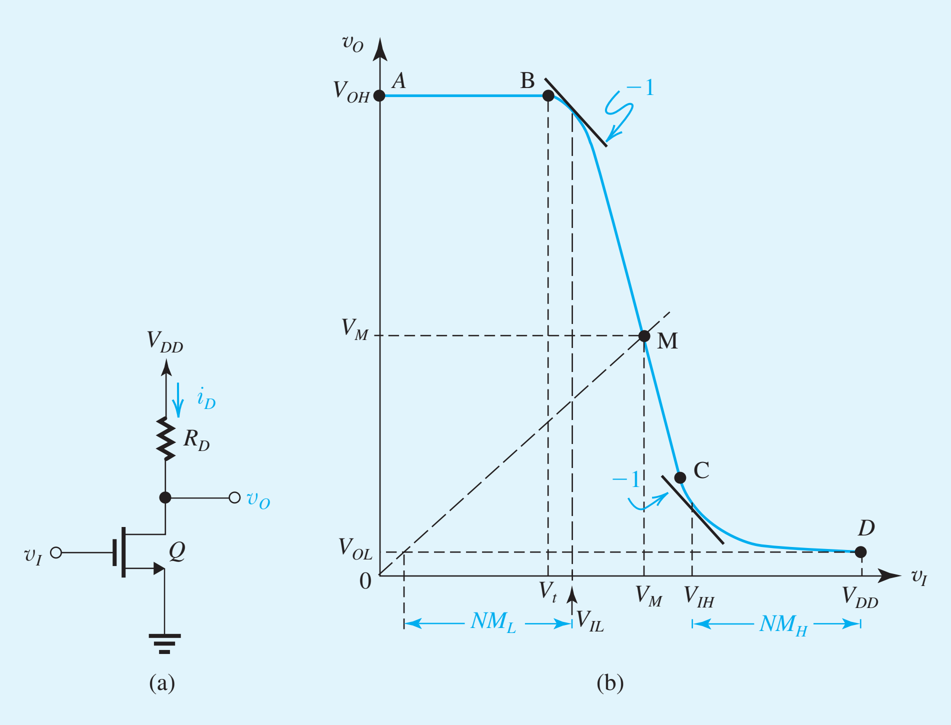


Figure 14.20 The resistively loaded MOS inverter and its VTC (Example 14.2).

To determine V_{IL}, we differentiate Eq. (14.11) and set $dv_O/dv_I = -1$,

$$\frac{dv_O}{dv_I} = -\frac{1}{V_x}(v_I - V_t)$$

$$-1 = -\frac{1}{V_x}(V_{IL} - V_t)$$

which results in

$$V_{IL} = V_t + V_x \tag{14.12}$$

To determine the coordinates of the midpoint M, we substitute $v_O = v_I = V_M$ in Eq. (14.11), thus obtaining

$$V_{DD} - V_M = \frac{1}{2V_x}(V_M - V_t)^2 \tag{14.13}$$

which can be solved to obtain

$$V_M = V_t + \sqrt{2(V_{DD} - V_t)V_x + V_x^2} - V_x \tag{14.14}$$

The boundary of the saturation-region segment BC, point C, is determined by substituting $v_O = v_I - V_t$ in Eq. (14.11) and solving for v_O to obtain

$$V_{OC} = \sqrt{2V_{DD}V_x + V_x^2} - V_x \tag{14.15}$$

and

$$V_{IC} = V_t + \sqrt{2V_{DD}V_x + V_x^2} - V_x \tag{14.16}$$

Beyond point C, the transistor operates in the triode region, thus

$$i_D = k_n\left[(v_I - V_t)v_O - \frac{1}{2}v_O^2\right]$$

and the output voltage is obtained as

$$v_O = V_{DD} - \frac{1}{V_x}\left[(v_I - V_t)v_O - \frac{1}{2}v_O^2\right] \tag{14.17}$$

which describes the segment CD of the VTC. To determine V_{IH}, we differentiate Eq. (14.17) and set $dv_O/dv_I = -1$:

$$\frac{dv_O}{dv_I} = -\left(\frac{1}{V_x}\right)\left[(v_I - V_t)\frac{dv_O}{dv_I} + v_O - v_O\frac{dv_O}{dv_I}\right]$$

$$-1 = -\frac{1}{V_x}[-(V_{IH} - V_t) + 2v_O]$$

which results in

$$V_{IH} - V_t = 2v_O - V_x \tag{14.18}$$

Substituting in Eq. (14.17) for v_I with the value of V_{IH} from Eq. (14.18) results in an equation in the value of v_O corresponding to $v_I = V_{IH}$, which can be solved to yield

$$v_O\big|_{v_I = V_{IH}} = 0.816\sqrt{V_{DD}V_x} \tag{14.19}$$

which can be substituted in Eq. (14.18) to obtain

$$V_{IH} = V_t + 1.63\sqrt{V_{DD}V_x} - V_x \tag{14.20}$$

Example 14.2 *continued*

To determine V_{OL} we substitute $v_I = V_{OH} = V_{DD}$ in Eq. (14.17):

$$V_{OL} = V_{DD} - \frac{1}{V_x}\left[(V_{DD} - V_t)V_{OL} - \frac{1}{2}V_{OL}^2\right] \tag{14.21}$$

Since we expect V_{OL} to be much smaller than $2(V_{DD} - V_t)$, we can approximate Eq. (14.21) as

$$V_{OL} \simeq V_{DD} - \frac{1}{V_x}(V_{DD} - V_t)V_{OL}$$

which results in

$$V_{OL} = \frac{V_{DD}}{1 + [(V_{DD} - V_t)/V_x]} \tag{14.22}$$

It is interesting to note that the value of V_{OL} can alternatively be found by noting that at point D, the MOSFET switch has a closure resistance r_{DS},

$$r_{DS} = \frac{1}{k_n(V_{DD} - V_t)} \tag{14.23}$$

and V_{OL} can be obtained from the voltage divider formed by R_D and r_{DS},

$$V_{OL} = V_{DD}\frac{r_{DS}}{R_D + r_{DS}} = \frac{V_{DD}}{1 + R_D/r_{DS}} \tag{14.24}$$

Substituting for r_{DS} from Eq. (14.23) gives an expression for V_{OL} identical to that in Eq. (14.22).

(b) We observe that all the inverter parameters derived above are functions of V_{DD}, V_t, and V_x only. Since V_{DD} and V_t are determined by the process technology, the only design parameter available is $V_x \equiv 1/k_n R_D$. To place V_M at half the supply voltage V_{DD}, we substitute $V_M = V_{DD}/2$ in Eq. (14.13) to obtain the value V_x must have as

$$V_x\big|_{V_M = V_{DD}/2} = \frac{(V_{DD}/2 - V_t)^2}{V_{DD}} \tag{14.25}$$

(c) For $V_{DD} = 1.8$ V and $V_t = 0.5$, we use Eq. (14.25) to obtain

$$V_x\big|_{V_M = 0.9\text{ V}} = \frac{(1.8/2 - 0.5)^2}{1.8} = 0.089\text{ V}$$

From Eq. (14.10): $V_{OH} = 1.8$ V

From Eq. (14.22): $V_{OL} = 0.12$ V

From Eq. (14.12): $V_{IL} = 0.59$ V

From Eq. (14.20): $V_{IH} = 1.06$ V

$$NM_L = V_{IL} - V_{OL} = 0.47\text{ V}$$
$$NM_H = V_{OH} - V_{IH} = 0.74\text{ V}$$

(d) To determine R_D, we use

$$k_n R_D = \frac{1}{V_x} = \frac{1}{0.089} = 11.24$$

Thus,

$$R_D = \frac{11.24}{k_n'(W/L)} = \frac{11.24}{300 \times 10^{-6} \times 1.5} = 25\text{ k}\Omega$$

The inverter dissipates power only when the output is low, in which case the current drawn from the supply is

$$I_{DD} = \frac{V_{DD} - V_{OL}}{R_D} = \frac{1.8 - 0.12}{25\text{ k}\Omega} = 67\ \mu\text{A}$$

and the power drawn from the supply during the low-output interval is

$$P_D = V_{DD} I_{DD} = 1.8 \times 67 = 121\ \mu\text{W}$$

Since the inverter spends half of the time in this state,

$$P_{D\text{average}} = \frac{1}{2} P_D = 60.5\ \mu\text{W}$$

(e) We now can make a few comments on the characteristics of this inverter circuit in comparison to the ideal characteristics:

1. The output signal swing, though not equal to the full power supply, is reasonably good: $V_{OH} = 1.8$ V, $V_{OL} = 0.12$ V.
2. The noise margins, though of reasonable values, are far from the optimum value of $V_{DD}/2$. This is particularly the case for NM_L.
3. Most seriously, the gate dissipates a relatively large amount of power. To appreciate this point, consider an IC chip with a million inverters (a small number by today's standards): Its power dissipation will be 61 W. This is too large, especially given that this is "static power," unrelated to the switching activity of the gates (more on this later).

We consider this inverter implementation to be entirely unsuitable for IC fabrication because each inverter requires a load resistance of 25 kΩ, a value that needs a large chip area (see Appendix A). To overcome this problem, we investigate in Example 14.3 the replacement of the passive resistance R_D with a PMOS transistor.

EXERCISES

D14.3 In an attempt to reduce the required value of R_D to 10 kΩ, the designer of the inverter in Example 14.2 decides to keep the parameter V_x unchanged but increases W/L. What is the new value required for W/L? Do the noise margins change? What does the power dissipation become?
Ans. 3.75; no; 151 μW

D14.4 In an attempt to reduce the required value of R_D to 10 kΩ, the designer of the inverter in Example 14.2 decides to change V_x while keeping W/L unchanged. What new value of V_x is needed? What do the noise margins become? What does the power dissipation become?
Ans. $V_x = 0.22$ V; $NM_L = 0.46$ V, $NM_H = 0.49$ V; 139 μW

Example 14.3 The Pseudo-NMOS Inverter

To eliminate the problem associated with the need for a large resistance R_D in the circuit of Fig. 14.20(a), studied in Example 14.2, R_D can be replaced by a MOSFET. One such possibility is the circuit in Fig. 14.21, where the load is a PMOS transistor Q_P whose gate is tied to ground in order to turn it on. Because of its resemblance to an earlier form of logic (NMOS logic, now obsolete) in which the load is an NMOS transistor, this circuit is known as a pseudo-NMOS inverter.

(a) Assuming $\lambda_1 = \lambda_2 = 0$, $V_{tn} = -V_{tp} = V_t$, and $k_n = 5k_p$, find V_{OH} and V_{OL}.
(b) For $k_n = 300\ \mu\text{A/V}^2$, $V_t = 0.4$ V, and $V_{DD} = 1.8$ V, evaluate the values of V_{OH} and V_{OL} and find the average power dissipated in the inverter, assuming it spends half the time in each of its two states.

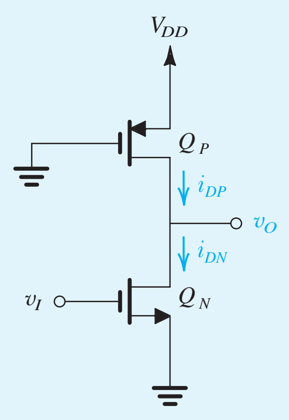


Figure 14.21 Pseudo-NMOS inverter for Example 14.3.

Solution

(a) To find V_{OH}, we set $v_I = 0$. Clearly Q_N will be off and conducting zero current. Transistor Q_P also will be conducting zero current but because its $V_{SG} = V_{DD}$, it will be operating in the triode region with a zero voltage between its source and drain; thus the output voltage will be equal to V_{DD},

$$V_{OH} = V_{DD}$$

Next, we find V_{OL} by setting $v_I = V_{DD}$. Transistor Q_N will be conducting. Since the output voltage V_{OL} will likely be low and thus lower than V_t, Q_P will be operating in the saturation region, thus

$$i_{DP} = \frac{1}{2} k_p (V_{DD} - V_t)^2$$

Q_N will be operating in the triode region, thus

$$i_{DN} = k_n \left[(V_{DD} - V_t)V_{OL} - \frac{1}{2} V_{OL}^2 \right]$$

Equating i_{DP} and i_{DN} yields a quadratic equation in V_{OL} that can be solved to obtain

$$V_{OL} = (V_{DD} - V_t)\left[1 - \sqrt{1 - (k_p/k_n)}\right]$$

Here we have rejected the other root of the quadratic equation on the assumption that its value will be greater than V_t and thus contravening our original assumption. The numerical values can be used to check these assumptions.

(b) Substituting the given numerical values we obtain

$$V_{OH} = 1.8 \text{ V}$$
$$V_{OL} = 0.15 \text{ V}$$

We note that V_{OL} is indeed lower than V_t, as originally assumed.

The inverter dissipates power in only one of its two states; namely, when its output is low. In this case, Q_P is operating in saturation and

$$\begin{aligned} i_{DP} &= \frac{1}{2} k_p (V_{DD} - V_t)^2 \\ &= \frac{1}{2} \times \left(\frac{300}{5}\right)(1.8 - 0.4)^2 \\ &= 58.8 \ \mu\text{A} \end{aligned}$$

and the power dissipation can be found from

$$P = i_{DP} V_{DD} = 58.8 \times 1.8 = 105.8 \ \mu\text{W}$$

The average power dissipation can now be found as

$$P_{av} = \frac{1}{2} \times 105.8 = 52.9 \ \mu\text{W}$$

EXERCISE

14.5 It is required to find V_M for the pseudo-NMOS inverter of Fig. 14.21. Recall that V_M is defined as a value of v_I that results in $v_O = V_M$. Convince yourself that because $V_M > V_t$, Q_N will be operating in saturation and Q_P will be operating in the triode region. Hence show that

$$V_M = V_t + \frac{V_{DD} - V_t}{\sqrt{r+1}}$$

when $r \equiv k_n/k_P$.
Evaluate V_M for $V_{DD} = 1.8$ V, $V_t = 0.4$ V, and $r = 5$.
Ans. 0.97 V

14.3 The CMOS Inverter

In this section we study the inverter circuit of the most widely used digital IC technology: CMOS. The basic CMOS inverter, synthesized in Section 14.1.2, is shown in Fig. 14.22. It utilizes two MOSFETs: one, Q_N, with an n channel and the other, Q_P, with a p channel. The body of each device is connected to its source, and thus no body effect arises. As will be seen shortly, the CMOS circuit realizes the conceptual inverter implementation studied in the previous section (Fig. 14.18), where a pair of switches are operated in a complementary fashion by the input voltage v_I.

14.3.1 Circuit Operation

We first consider the two extreme cases: when v_I is at logic-0 level, which is 0 V, and when v_I is at logic-1 level, which is V_{DD} volts. In both cases, for ease of exposition we shall consider the n-channel device Q_N to be the driving transistor and the p-channel device Q_P to be the

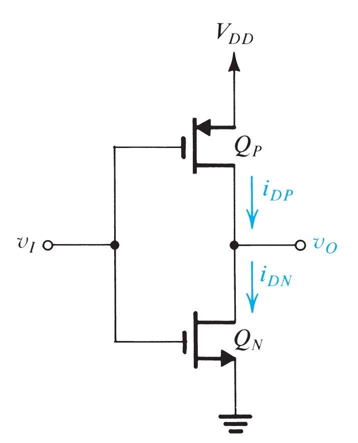


Figure 14.22 The CMOS inverter.

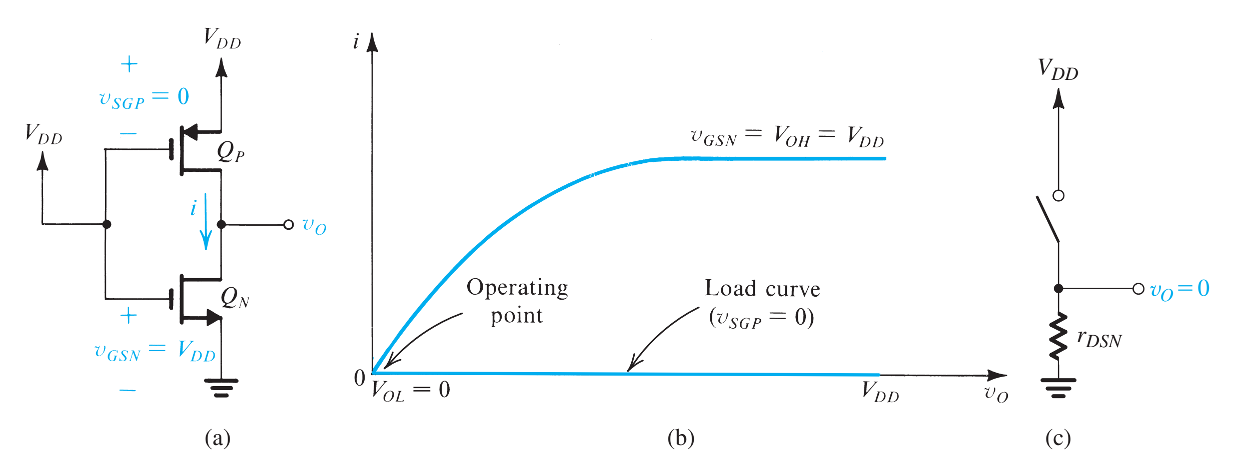


Figure 14.23 Operation of the CMOS inverter when v_I is high: **(a)** circuit with $v_I = V_{DD}$ (logic-1 level, or V_{OH}); **(b)** graphical construction to determine the operating point; **(c)** equivalent circuit.

load. However, since the circuit is symmetric, this assumption is obviously arbitrary, and the reverse would lead to identical results.

Figure 14.23 illustrates the case when $v_I = V_{DD}$, showing the $i_D - v_{DS}$ characteristic curve for Q_N with $v_{GSN} = V_{DD}$. (Note that $i_D = i$ and $v_{DSN} = v_O$.) Superimposed on the Q_N characteristic curve is the load curve, which is the $i_D - v_{SD}$ curve of Q_P for the case $v_{SGP} = 0$ V. Since $v_{SGP} < |V_t|$, the load curve will be a horizontal straight line at zero current level. The operating point will be at the intersection of the two curves, where we note that the output voltage is zero and the current through the two devices is also zero. This means that the power dissipation in the circuit is zero. Note, however, that although Q_N is operating at zero current and zero drain-source voltage (i.e., at the origin of the $i_D - v_{DS}$ plane), the operating point is on a steep segment of the $i_D - v_{DS}$ characteristic curve. Thus Q_N provides a low-resistance path between the output terminal and ground, with the resistance obtained using Eq. (5.13b) as

$$r_{DSN} = 1 \Big/ \left[k_n' \left(\frac{W}{L} \right)_n (V_{DD} - V_{tn}) \right] \tag{14.26}$$

Figure 14.23(c) shows the equivalent circuit of the inverter when the input is high.[2] This circuit confirms that $v_O \equiv V_{OL} = 0$ V and that the power dissipation in the inverter is zero.

The other extreme case, when $v_I = 0$ V, is illustrated in Fig. 14.24. In this case Q_N is operating at $v_{GSN} = 0$; hence its $i_D - v_{DS}$ characteristic is a horizontal straight line at zero current level. The load curve is the $i_D - v_{SD}$ characteristic of the p-channel device with $v_{SGP} = V_{DD}$. As shown, at the operating point the output voltage is equal to V_{DD}, and the current in the two devices is still zero. Thus the power dissipation in the circuit is zero in both extreme states.

[2] In Section 14.1 we referred to r_{DSN} (and r_{DSP} for p-channel devices) as R_{on}.

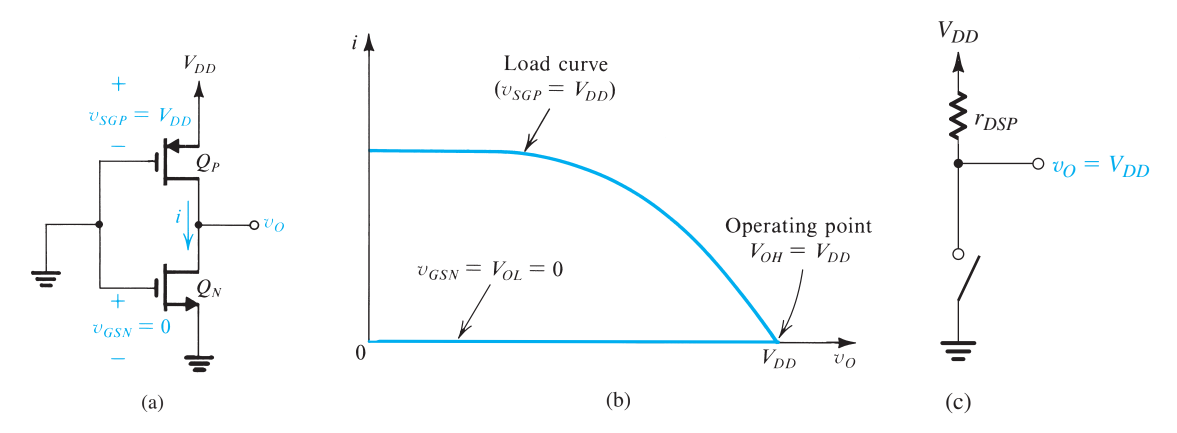
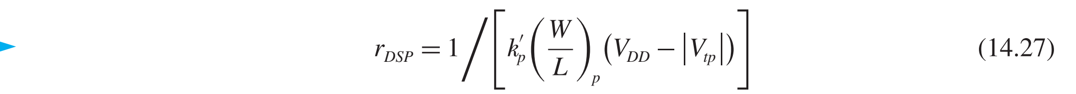


Figure 14.24 Operation of the CMOS inverter when v_I is low: **(a)** circuit with $v_I = 0$ V (logic-0 level, or V_{OL}); **(b)** graphical construction to determine the operating point; **(c)** equivalent circuit.

Figure 14.24(c) shows the equivalent circuit of the inverter when the input is low. Here we see that Q_P provides a low-resistance path between the output terminal and the dc supply V_{DD}, with the resistance given by



$$r_{DSP} = 1 \Big/ \left[k_p' \left(\frac{W}{L} \right)_p (V_{DD} - |V_{tp}|) \right] \tag{14.27}$$

The equivalent circuit confirms that in this case $v_O \equiv V_{OH} = V_{DD}$ and that the power dissipation in the inverter is zero.

It should be noted, however, that in spite of the fact that the quiescent current is zero, the load-driving capability of the CMOS inverter is high. For instance, with the input high, as in the circuit of Fig. 14.23, transistor Q_N can sink a relatively large load current. This current can quickly discharge the load capacitance, as will be seen shortly. Because of its action in sinking load current and thus pulling the output voltage down toward ground, transistor Q_N is known as the pull-down device. Similarly, with the input low, as in the circuit of Fig. 14.24, transistor Q_P can source a relatively large load current. This current can quickly charge up a load capacitance, thus pulling the output voltage up toward V_{DD}. Hence, Q_P is known as the pull-up device. The reader will recall that we used this terminology in connection with the conceptual inverter circuit of Fig. 14.18 and in Section 14.1 as well.

From the above, we conclude that the basic CMOS logic inverter behaves as an ideal inverter. In summary:

1. The output voltage levels are 0 and V_{DD}, and thus the signal swing is the maximum possible. This, coupled with the fact that the inverter can be designed to provide a symmetrical voltage-transfer characteristic, results in wide noise margins.
2. The static power dissipation in the inverter is zero (neglecting the dissipation due to leakage currents) in both of its states. This is because no dc path exists between the power supply and ground in either state.
3. A low-resistance path exists between the output terminal and ground (in the low-output state) or V_{DD} (in the high-output state). These low-resistance paths ensure that the output voltage is 0 or V_{DD} independent of the exact values of the W/L ratios or

other device parameters. Furthermore, the low output resistance makes the inverter less sensitive to the effects of noise and other disturbances.

4. The active pull-up and pull-down devices provide the inverter with high output-driving capability in both directions. As will be seen in Section 14.4, this speeds up the operation considerably.
5. The input resistance of the inverter is infinite (because $I_G = 0$). Thus the inverter can drive an arbitrarily large number of similar inverters with no loss in signal level. Of course, each additional inverter increases the load capacitance on the driving inverter and slows down the operation. In Section 14.4, we will consider the inverter switching times.

FRANK MARION WANLESS—THE INVENTOR OF CMOS:

While working for Fairchild Semiconductor in 1963, Frank Wanless filed the first patent on CMOS logic, heralding the new age of zero-static-power logic. In 1964, as director of research and engineering at General Microelectronics (a start-up later bought by Philco-Ford), he created the first commercial CMOS integrated circuit. The symmetry of the logic form Wanless had invented was at first emphasized by the use of the name COmplementary Symmetry MOS, or COS-MOS, but the simpler CMOS shorthand soon prevailed.

14.3.2 The Voltage-Transfer Characteristic (VTC)

The complete voltage-transfer characteristic (VTC) of the CMOS inverter can be obtained by repeating the graphical procedure, used above in the two extreme cases, for all intermediate values of v_I. In the following, we shall calculate the critical points of the resulting voltage-transfer curve. For this we need the i–v relationships of Q_N and Q_P. For Q_N,

$$i_{DN} = k_n'\left(\frac{W}{L}\right)_n\left[(v_I - V_{tn})v_O - \frac{1}{2}v_O^2\right] \quad \text{for } v_O \le v_I - V_{tn} \tag{14.28}$$

and

$$i_{DN} = \frac{1}{2}k_n'\left(\frac{W}{L}\right)_n(v_I - V_{tn})^2 \quad \text{for } v_O \ge v_I - V_{tn} \tag{14.29}$$

For Q_P,

$$i_{DP} = k_p'\left(\frac{W}{L}\right)_p\left[(V_{DD} - v_I - |V_{tp}|)(V_{DD} - v_O) - \frac{1}{2}(V_{DD} - v_O)^2\right] \quad \text{for } v_O \ge v_I + |V_{tp}| \tag{14.30}$$

and

$$i_{DP} = \frac{1}{2}k_p'\left(\frac{W}{L}\right)_p(V_{DD} - v_I - |V_{tp}|)^2 \quad \text{for } v_O \le v_I + |V_{tp}| \tag{14.31}$$

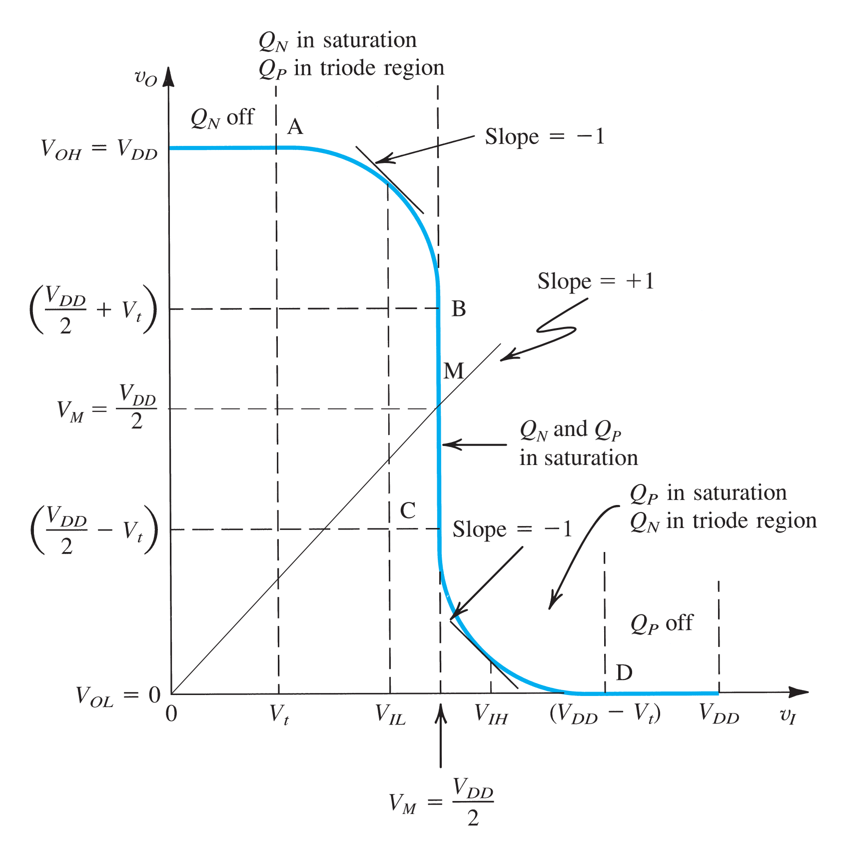


Figure 14.25 The voltage-transfer characteristic of the CMOS inverter when Q_N and Q_P are matched.

The CMOS inverter is usually designed to have $V_{tn} = |V_{tp}| = V_t$. Also, although this is not always the case, we shall assume that Q_N and Q_P are matched; that is, $k_n'(W/L)_n = k_p'(W/L)_p$. It should be noted that since μ_p is 0.25 to 0.5 times the value of μ_n, to make $k'(W/L)$ of the two devices equal, the width of the p-channel device is made two to four times that of the n-channel device. More specifically, the two devices are designed to have equal lengths, with widths related by

$$\frac{W_p}{W_n} = \frac{\mu_n}{\mu_p} \tag{14.32}$$

This will result in $k_n'(W/L)_n = k_p'(W/L)_p$, and the inverter will have a symmetric transfer characteristic and equal current-driving capability in both directions (pull-up and pull-down).

With Q_N and Q_P matched, the CMOS inverter has the voltage-transfer characteristic shown in Fig. 14.25. As indicated, the transfer characteristic has five distinct segments corresponding to different combinations of modes of operation of Q_N and Q_P. The vertical segment BC is obtained when both Q_N and Q_P are operating in the saturation region. Because we are neglecting the finite output resistance in saturation, that is, assuming $\lambda_N = \lambda_P = 0$, the inverter gain in this region is infinite. From symmetry, this vertical segment occurs at $v_I = V_{DD}/2$ and is bounded by $v_O(\text{B}) = V_{DD}/2 + V_t$, at which value Q_P enters the triode region and $v_O(\text{C}) = V_{DD}/2 - V_t$, at which value Q_N enters the triode region.

The reader will recall from Section 14.2.1 that in addition to V_{OL} and V_{OH}, two other points on the transfer curve determine the noise margins of the inverter. These are the maximum permitted logic-0 or "low" level at the input, V_{IL}, and the minimum permitted logic-1 or "high" level at the input, V_{IH}. These are formally defined as the two points on the transfer curve at which the incremental gain is unity (i.e., the slope is –1 V/V).

To determine V_{IH}, we note that Q_N is in the triode region, and thus its current is given by Eq. (14.28), while Q_P is in saturation and its current is given by Eq. (14.31). Equating i_{DN} and i_{DP}, and assuming matched devices, gives

$$(v_I - V_t)v_O - \frac{1}{2}v_O^2 = \frac{1}{2}(V_{DD} - v_I - V_t)^2 \tag{14.33}$$

Differentiating both sides relative to v_I results in

$$(v_I - V_t)\frac{dv_O}{dv_I} + v_O - v_O\frac{dv_O}{dv_I} = -(V_{DD} - v_I - V_t)$$

in which we substitute $v_I = V_{IH}$ and $dv_O/dv_I = -1$ to obtain

$$v_O = V_{IH} - \frac{V_{DD}}{2} \tag{14.34}$$

Substituting $v_I = V_{IH}$ and for v_O from Eq. (14.34) in Eq. (14.33) gives

$$V_{IH} = \frac{1}{8}(5V_{DD} - 2V_t) \tag{14.35}$$

V_{IL} can be determined in a manner similar to that used to find V_{IH}. Alternatively, we can use the symmetry relationship

$$V_{IH} - \frac{V_{DD}}{2} = \frac{V_{DD}}{2} - V_{IL}$$

together with V_{IH} from Eq. (14.35) to obtain

$$V_{IL} = \frac{1}{8}(3V_{DD} + 2V_t) \tag{14.36}$$

The noise margins can now be determined as follows:

$$\begin{aligned} NM_H &= V_{OH} - V_{IH} \\ &= V_{DD} - \frac{1}{8}(5V_{DD} - 2V_t) \\ &= \frac{1}{8}(3V_{DD} + 2V_t) \end{aligned} \tag{14.37}$$

$$\begin{aligned} NM_L &= V_{IL} - V_{OL} \\ &= \frac{1}{8}(3V_{DD} + 2V_t) - 0 \\ &= \frac{1}{8}(3V_{DD} + 2V_t) \end{aligned} \tag{14.38}$$

As expected, the symmetry of the voltage-transfer characteristic results in equal noise margins. Of course, if Q_N and Q_P are not matched, the voltage-transfer characteristic will no longer be symmetric, and the noise margins will not be equal.

14.3.3 The Situation When Q_N and Q_P Are Not Matched

In the above we assumed that Q_N and Q_P are matched; that is, in addition to $V_{tn} = |V_{tp}|$, the transconductance parameters k_n and k_p are made equal by selecting W_p/W_n according to Eq. (14.32). The result is a symmetrical VTC that switches at the midpoint of the supply; that is, $V_M = V_{DD}/2$. The symmetry, as we have seen, equalizes and maximizes the noise margins.

The price paid for obtaining a perfectly symmetric VTC is that the width of the p-channel device can be three to four times as large as that of the n-channel device. This can result in a relatively large silicon area, which, besides being wasteful of silicon real estate, can also result in increased device capacitances and a corresponding increase in the propagation delay of the inverter (Section 14.4). It is useful, therefore, to inquire into the effect of not matching Q_N and Q_P. Toward that end we derive an expression for the switching voltage V_M as follows.

Since at M, both Q_N and Q_P operate in saturation, their currents are given by Eqs. (14.29) and (14.31), respectively. Substituting $v_I = v_O = V_M$, and equating the two currents results in

$$V_M = \frac{r\left(V_{DD} - |V_{tp}|\right) + V_{tn}}{r+1} \tag{14.39}$$

where

$$r = \sqrt{\frac{k_p}{k_n}} = \sqrt{\frac{\mu_p W_p}{\mu_n W_n}} \tag{14.40}$$

where we have assumed that Q_N and Q_P have the same channel length L, which is usually the case with L equal to the minimum available for the given process technology. Note that the matched case corresponds to $r = 1$. For $|V_{tp}| = V_{tn}$, and $r = 1$, Eq. (14.39) yields $V_M = V_{DD}/2$, as expected. For a given process, that is, given values for V_{DD}, V_{tn}, and V_{tp}, one can plot V_M versus the matching parameter r. Such a plot, for a 0.18-μm process, is shown in Fig. 14.26. We make the following two observations:

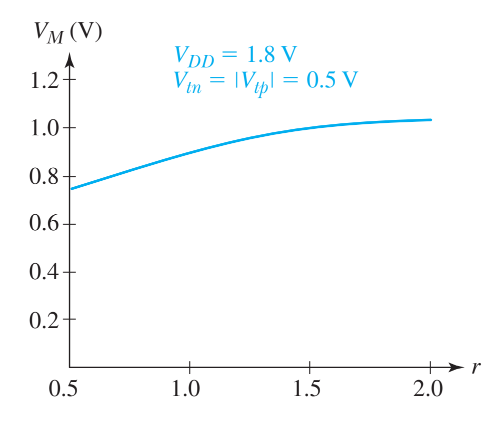


Figure 14.26 Variation of the inverter switching voltage, V_M, with the parameter $r = \sqrt{k_p/k_n}$.

1. V_M increases with r. Thus, making $k_p > k_n$ shifts V_M toward V_{DD}. Conversely, making $k_p < k_n$ shifts V_M toward 0.
2. V_M is not a strong function of r. For the particular case shown, lowering r by a factor of 2 (from 1 to 0.5), reduces V_M by only 0.13 V.

Observation 2 implies that if one is willing to tolerate a small reduction in NM_L, substantial savings in silicon area can be obtained. This point is illustrated in Example 14.4.

Example 14.4 CMOS Inverter Static Characteristics and Design

Consider a CMOS inverter fabricated in a 0.18-μm process for which $V_{DD} = 1.8$ V, $V_{tn} = |V_{tp}| = 0.5$ V, $\mu_n = 4\mu_p$, and $\mu_n C_{ox} = 300\ \mu\text{A/V}^2$. In addition, Q_N and Q_P have $L = 0.18\ \mu$m and $(W/L)_n = 1.5$.

(a) Find W_p that results in $V_M = V_{DD}/2 = 0.9$ V. What is the silicon area utilized by the inverter in this case?

(b) For the matched case in (a), find the values of V_{OH}, V_{OL}, V_{IH}, V_{IL}, and the noise margins NM_L and NM_H. For $v_I = V_{IH}$, what value of v_O results? This can be considered the worst-case value of V_{OL}. Similarly, for $v_I = V_{IL}$, find v_O that is the worst-case value of V_{OH}. Now, use these worst-case values to determine more conservative values for the noise margins.

(c) For the matched case in (a), find the output resistance of the inverter in each of its two states.

(d) If $\lambda_n = |\lambda_p| = 0.2\ \text{V}^{-1}$, what is the inverter gain at $v_I = V_M$? If a straight line is drawn through the point $v_I = v_O = V_M$ with a slope equal to the gain, at what values of v_I does it intercept the horizontal lines $v_O = 0$ and $v_O = V_{DD}$? Use these intercepts to estimate the width of the transition region of the VTC.

(e) If $W_p = W_n$, what value of V_M results? What do you estimate the reduction of NM_L (relative to the matched case) to be? What is the percentage savings in silicon area (relative to the matched case)?

(f) Repeat (e) for the case $W_p = 2W_n$. This case, which is frequently used in industry, can be considered to be a compromise between the minimum-area case in (e) and the matched case.

Solution

(a) To obtain $V_M = V_{DD}/2 = 0.9$ V, we select W_p according to Eq. (14.32),

$$\frac{W_p}{W_n} = \frac{\mu_n}{\mu_p} = 4$$

Since $W_n/L = 1.5$, $W_n = 1.5 \times 0.18 = 0.27\ \mu$m. Thus,

$$W_p = 4 \times 0.27 = 1.08\ \mu\text{m}$$

For this design, the silicon area is

$$A = W_n L + W_p L = L(W_n + W_p)$$
$$= 0.18(0.27 + 1.08) = 0.243\ \mu\text{m}^2$$

(b)

$$V_{OH} = V_{DD} = 1.8\ \text{V}$$

$$V_{OL} = 0\ \text{V}$$

Example 14.4 *continued*

To obtain V_{IH} we use Eq. (14.35),

$$V_{IH} = \frac{1}{8}(5V_{DD} - 2V_t) = \frac{1}{8}(5 \times 1.8 - 2 \times 0.5) = 1\text{ V}$$

To obtain V_{IL} we use Eq. (14.36),

$$V_{IL} = \frac{1}{8}(3V_{DD} + 2V_t) = \frac{1}{8}(3 \times 1.8 + 2 \times 0.5) = 0.8\text{ V}$$

We can now compute the noise margins as

$$NM_H = V_{OH} - V_{IH} = 1.8 - 1.0 = 0.8\text{ V}$$

$$NM_L = V_{IL} - V_{OL} = 0.8 - 0 = 0.8\text{ V}$$

As expected, $NM_H = NM_L$, and their value is very close to the optimum value of $V_{DD}/2 = 0.9$ V.

For $v_I = V_{IH} = 1$ V, we can obtain the corresponding value of v_O by substituting in Eq. (14.34),

$$v_O = V_{IH} - \frac{V_{DD}}{2} = 1 - \frac{1.8}{2} = 0.1\text{ V}$$

Thus, the worst-case value of V_{OL}, that is, $V_{OL\max}$, is 0.1 V, and the noise margin NM_L reduces to

$$NM_L = V_{IL} - V_{OL\max} = 0.8 - 0.1 = 0.7\text{ V}$$

From symmetry, we can obtain the value of v_O corresponding to $v_I = V_{IL}$ as

$$v_O = V_{DD} - 0.1 = 1.7\text{ V}$$

Thus the worst-case value of V_{OH}, that is, $V_{OH\min}$, is 1.7 V, and the noise margin NM_H reduces to

$$NM_H = V_{OH\min} - V_{IH} = 1.7 - 1 = 0.7\text{ V}$$

Note that the reduction in the noise margins is slight.

(c) The output resistance of the inverter in the low-output state is

$$r_{DSN} = \frac{1}{\mu_n C_{ox}(W/L)_n(V_{DD} - V_{tn})}$$

$$= \frac{1}{300 \times 10^{-6} \times 1.5(1.8 - 0.5)} = 1.71\text{ k}\Omega$$

Since Q_N and Q_P are matched, the output resistance in the high-output state will be equal, that is,

$$r_{DSP} = r_{DSN} = 1.71 \text{ k}\Omega$$

(d) If the inverter is biased to operate at $v_I = v_O = V_M = 0.9$ V, then each of Q_N and Q_P will be operating at an overdrive voltage $V_{OV} = V_M - V_t = 0.9 - 0.5 = 0.4$ V and will be conducting equal dc currents I_D of

$$\begin{aligned} I_D &= \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_N V_{OV}^2 \\ &= \frac{1}{2} \times 300 \times 1.5 \times 0.4^2 \\ &= 36\ \mu\text{A} \end{aligned}$$

Thus, Q_N and Q_P will have equal transconductances:

$$g_{mn} = g_{mp} = \frac{2I_D}{V_{OV}} = \frac{2 \times 36}{0.4} = 0.18 \text{ mA/V}^2$$

Transistors Q_N and Q_P will have equal output resistances r_o,

$$r_{on} = r_{op} = \frac{|V_A|}{I_D} = \frac{1}{|\lambda| I_D} = \frac{1}{0.2 \times 36} = 139 \text{ k}\Omega$$

We can now compute the voltage gain at M as

$$\begin{aligned} A_v &= -(g_{mn} + g_{mp})(r_{on} \,\|\, r_{op}) \\ &= -(0.18 + 0.18)(139 \,\|\, 139) = -25 \text{ V/V} \end{aligned}$$

When the straight line at M of slope −25 V/V is extrapolated, it intersects the line $v_O = 0$ at $[0.9 + 0.9/25] = 0.936$ V and the line $v_O = V_{DD}$ at $0.9 - 0.9/25 = 0.864$ V. Thus the width of the transition region can be considered to be $(0.936 - 0.864) = 0.072$ V.

(e) For $W_p = W_n$, the parameter r can be found from Eq. (14.40),

$$r = \sqrt{\frac{\mu_p W_p}{\mu_n W_n}} = \sqrt{\frac{1}{4} \times 1} = 0.5$$

The corresponding value of V_M can be determined from Eq. (14.39) as

$$V_M = \frac{0.5(1.8 - 0.5) + 0.5}{0.5 + 1} = 0.77 \text{ V}$$

Example 14.4 *continued*

Thus V_M shifts by only -0.13 V. Without recalculating V_{IL} we can estimate the reduction in NM_L to be approximately equal to the shift in V_M, that is, NM_L becomes $0.8 - 0.13 = 0.67$ V. The silicon area for this design can be computed as follows:

$$A = L(W_n + W_p) = 0.18(0.27 + 0.27)$$
$$= 0.0972\ \mu\text{m}^2$$

This represents a 60% reduction from the matched case!

(f) For $W_p = 2W_n$,

$$r = \sqrt{\frac{1}{4} \times 2} = \frac{1}{\sqrt{2}} = 0.707$$

$$V_M = \frac{0.707(1.8 - 0.5) + 0.5}{0.707 + 1} = 0.83\ \text{V}$$

Thus, relative to the matched case, the shift in V_M is only -0.07 V. We estimate that NM_L will decrease from 0.8 V by the same amount; thus NM_L becomes 0.73 V. In this case, the silicon area required is

$$A = L(W_n + W_p) = 0.18(0.27 + 0.54)$$
$$= 0.146\ \mu\text{m}^2$$

which represents a 40% reduction relative to the matched case!

EXERCISES

14.6 Consider a CMOS inverter fabricated in a 0.13-μm process for which $V_{DD} = 1.2$ V, $V_{tn} = -V_{tp} = 0.4$ V, $\mu_n/\mu_p = 4$, and $\mu_n C_{ox} = 430\ \mu\text{A/V}^2$. In addition, Q_N and Q_P have $L = 0.13\ \mu$m and $(W/L)_n = 1.0$.

(a) Find W_p that results in $V_M = 0.6$ V.

(b) For the matched case in (a), find the values of V_{OH}, V_{OL}, V_{IH}, V_{IL}, NM_H, and NM_L.

(c) For the inverter in (a), find the output resistance in each of its two states.

(d) For a minimum-size inverter for which $(W/L)_p = (W/L)_n = 1.0$, find V_M.

Ans. (a) 0.52 μm; (b) 1.2 V, 0 V, 0.65 V, 0.55 V, 0.55 V, 0.55 V; (c) 2.9 kΩ, 2.9 kΩ; (d) 0.53 V

D14.7 A CMOS inverter utilizes $V_{DD} = 5$ V, $V_{tn} = |V_{tp}| = 1$ V, and $\mu_n C_{ox} = 2\mu_p C_{ox} = 50\ \mu\text{A/V}^2$. Find $(W/L)_n$ and $(W/L)_p$ so that $V_M = 2.5$ V and so that for $v_I = V_{DD}$, the inverter can sink a current of 0.2 mA with the output voltage not exceeding 0.2 V.

Ans. $(W/L)_n \simeq 5$; $(W/L)_p \simeq 10$

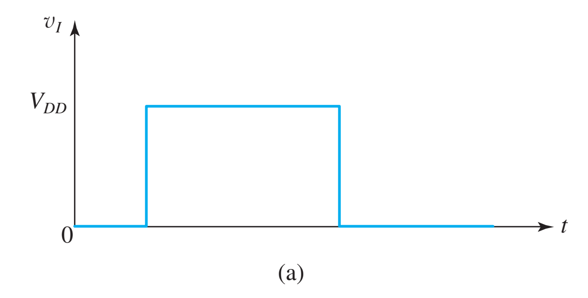
14.4 Dynamic Operation of the CMOS Inverter

The speed of operation of a digital system (e.g., a computer) is determined by the propagation delay of the logic gates used to construct the system. Since the inverter is the basic logic gate of any digital IC technology, the propagation delay of the inverter is a fundamental parameter in characterizing the speed of a given technology. We begin our study of the dynamic operation of CMOS in Section 14.4.1 by considering the propagation delay of a general inverter circuit. There, we introduce key definitions and analysis methods that are applied in the CMOS case in Sections 14.4.2 and 14.4.3.

14.4.1 Propagation Delay

The propagation delay is the time the inverter takes to respond to a change at its input. To be specific, let us consider an inverter fed with the ideal pulse shown in Fig. 14.27(a). The resulting output signal of the inverter is shown in Fig. 14.27(b). We make the following two observations.

1. The output signal is no longer an ideal pulse. Rather, it has rounded edges; that is, the pulse takes some time to fall to its low value and to rise to its high value. We speak of this as the pulse having finite fall and rise times. We will provide a precise definition of these shortly.
2. There is a time delay between each edge of the input pulse and the corresponding change in the output of the inverter. If we define the "switching point" of the output as the time at which the output pulse passes through the half-point of its excursion, then we can define the propagation delays of the inverter as indicated in Fig. 14.27(b). Note that there are two propagation delays, which are not necessarily equal: the propagation delay for the output going from high to low, t_{PHL}, and the propagation delay for the



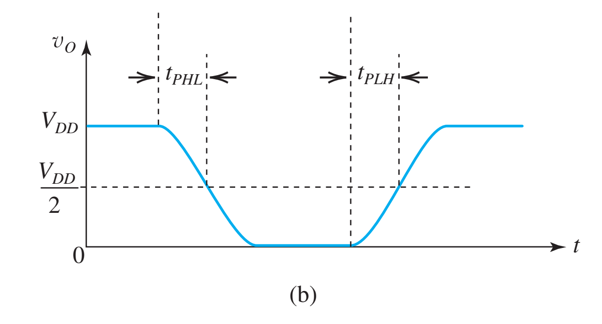


Figure 14.27 An inverter fed with the ideal pulse in **(a)** provides at its output the pulse in **(b)**. Two delay times are defined as indicated.

output going from low to high, t_{PLH}. The inverter propagation delay t_P is defined as the average of the two,

$$t_P \equiv \frac{1}{2}(t_{PLH} + t_{PHL}) \tag{14.41}$$

Having defined the inverter propagation delay, we now consider the *maximum* switching frequency of the inverter. From Fig. 14.27(b) we can see that the *minimum* period for each cycle is

$$T_{\min} = t_{PHL} + t_{PLH} = 2t_P \tag{14.42}$$

Thus the **maximum switching frequency**[3] is

$$f_{\max} = \frac{1}{T_{\min}} = \frac{1}{2t_P} \tag{14.43}$$

At this point the reader is no doubt wondering about the cause of the finite propagation time of the inverter. It is simply a result of the time needed to charge and discharge the various capacitances in the circuit. These include the MOSFET capacitances, the wiring capacitance, and the input capacitances of all the logic gates driven by the inverter. We will have a lot more to say about these capacitances and about the determination of t_P shortly. For the time being, however, we make two important points:

1. A fundamental relationship in analyzing the dynamic operation of a circuit is

$$I\Delta t = \Delta Q = C\Delta V \tag{14.44}$$

   That is, a current I flowing through a capacitance C for an interval Δt deposits a charge ΔQ on the capacitor, which causes the capacitor voltage to increase by ΔV.

2. A thorough familiarity with the time response of single-time-constant (STC) circuits is of great help in the analysis of the dynamic operation of digital circuits. A review of this subject is presented in Appendix E. For our purposes here, we remind the reader of the key equation in determining the response to a step function:

   Consider a step-function input applied to an STC circuit of either the low-pass or high-pass type, and let the circuit have a time constant τ. The output at any time t is given by

$$y(t) = Y_\infty - \left(Y_\infty - Y_{0+}\right)e^{-t/\tau} \tag{14.45}$$

   where Y_∞ is the final value, that is, the value toward which the response is heading, and Y_{0+} is the value of the response immediately after $t = 0$. This equation states that the output at any time t is equal to the difference between the final value Y_∞ and a gap whose initial value is $Y_\infty - Y_{0+}$ and that is shrinking exponentially.

[3]This is a theoretical upper bound; practical circuits are operated at frequencies 10 to 20 times lower.

Example 14.5 Calculating the Propagation Delay of a Simple Inverter

Return to the inverter of Fig. 14.17(a) and consider the case where a capacitor C is connected between the output node and ground. If at $t = 0$, v_I goes low, and assuming that the switch opens instantaneously, find the time for v_O to reach $\frac{1}{2}(V_{OH} + V_{OL})$. This is the low-to-high propagation time, t_{PLH}. Calculate the value of t_{PLH} for the case $R = 25\text{ k}\Omega$ and $C = 10\text{ fF}$.

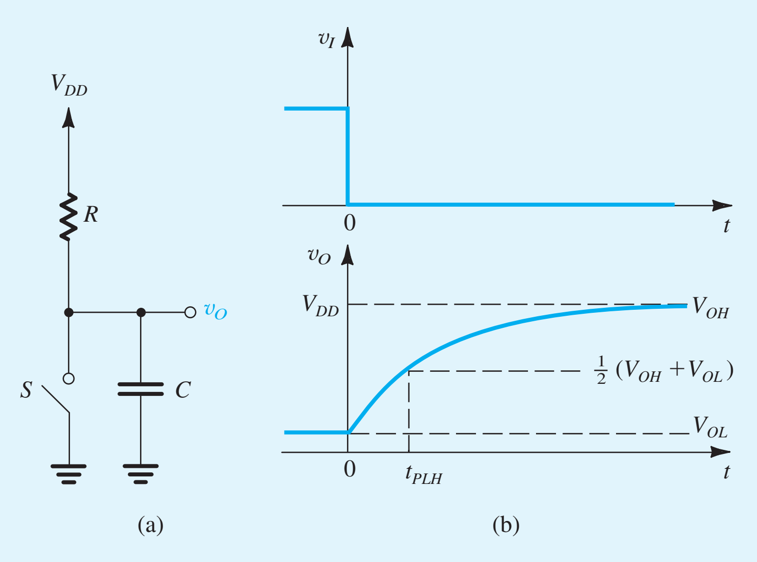


Figure 14.28 Example 14.5: **(a)** The inverter circuit after the switch opens (i.e., for $t \geq 0+$). **(b)** Waveforms of v_I and v_O. Observe that the switch is assumed to operate instantaneously. v_O rises exponentially, starting at V_{OL} and heading toward V_{OH}.

Solution

Before the switch opens, $v_O = V_{OL}$. When the switch opens at $t = 0$, the circuit takes the form shown in Fig. 14.28(a). Since the voltage across the capacitor cannot change instantaneously, at $t = 0+$ the output will still be V_{OL}. Then the capacitor charges through R, and v_O rises exponentially toward V_{DD}. The output waveform will be as shown in Fig. 14.28(b), and its equation can be obtained by substituting in Eq. (14.45): $v_O(\infty) = V_{OH} = V_{DD}$ and $v_O(0+) = V_{OL}$. Thus,

$$v_O(t) = V_{OH} - (V_{OH} - V_{OL})e^{-t/\tau}$$

where $\tau = CR$. To find t_{PLH}, we substitute

$$v_O(t_{PLH}) = \frac{1}{2}(V_{OH} + V_{OL})$$

Thus,

$$\frac{1}{2}(V_{OH} + V_{OL}) = V_{OH} - (V_{OH} - V_{OL})e^{-t_{PLH}/\tau}$$

which results in

$$t_{PLH} = \tau \ln 2 = 0.69\tau$$

Example 14.5 *continued*

Note that this expression is independent of the values of V_{OL} and V_{OH}. For the numerical values given,

$$t_{PLH} = 0.69RC$$
$$= 0.69 \times 25 \times 10^3 \times 10 \times 10^{-15}$$
$$= 173 \text{ ps}$$

EXERCISES

14.8 A capacitor C whose initial voltage is 0 is charged to a voltage V_{DD} by a constant-current source I. Find the time t_{PLH} at which the capacitor voltage reaches $(V_{DD}/2)$. What value of I is required to obtain a 10-ps propagation delay with $C = 10$ fF and $V_{DD} = 1.8$ V?
Ans. $t_{PLH} = CV_{DD}/2I$; 0.9 mA

14.9 For the inverter of Fig. 14.18(a), let the on-resistance of P_U be 20 kΩ and that of $P_D = 10$ kΩ. If the capacitance $C = 10$ fF, find t_{PLH}, t_{PHL}, and t_P.
Ans. 138 ps; 69 ps; 104 ps

We conclude this section by showing in Fig. 14.29 the formal definition of the propagation delay of an inverter. As shown, an input pulse with finite (nonzero) **rise and fall times** is applied. The inverted pulse at the output exhibits finite rise and fall times (labeled t_{TLH} and t_{THL}, where the subscript T denotes transition, LH denotes low to high, and HL denotes high to low). There is also a delay time between the input and output waveforms. The usual way to specify the propagation delay is to take the average of the high-to-low propagation delay, t_{PHL}, and the low-to-high propagation delay, t_{PLH}. As indicated, these delays are measured between the 50% points of the input and output waveforms. Also note that the **transition times** are specified using the 10% and 90% points of the output excursion ($V_{OH} - V_{OL}$).

EXERCISE

14.10 A capacitor $C = 100$ fF is discharged from a voltage V_{DD} to zero through a resistance $R = 2$ kΩ. Find the fall time t_f of the capacitor voltage.
Ans. $t_f \simeq 2.2CR = 0.44$ ns

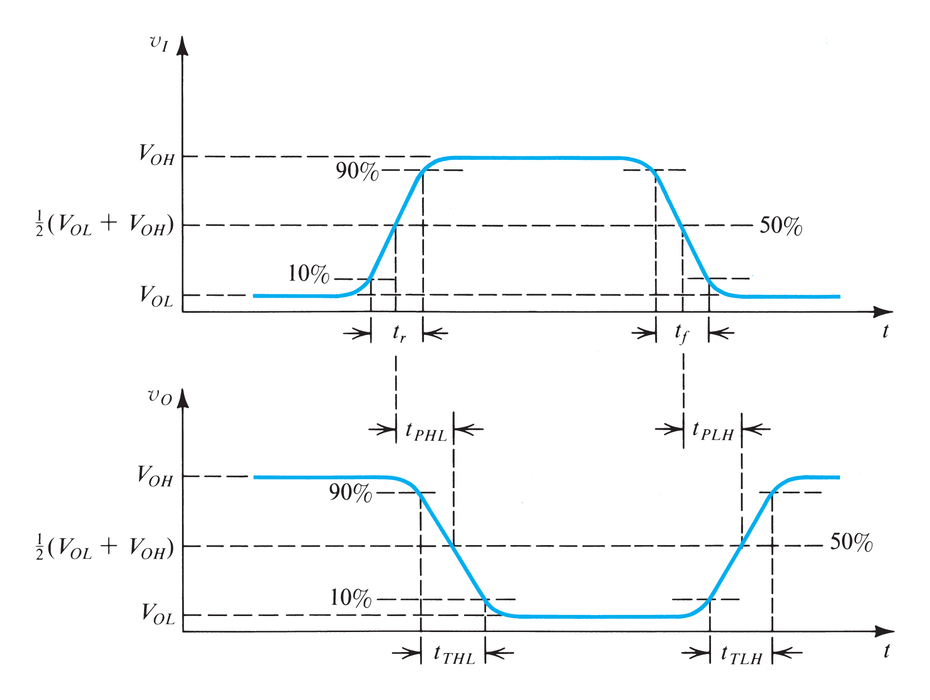


Figure 14.29 Definitions of propagation delays and transition times of the logic inverter.

14.4.2 Determining the Propagation Delay of the CMOS Inverter

Our strategy for determining the propagation delay of the CMOS inverter consists of two steps:

1. Replace all the capacitances in the circuit; that is, the various capacitances associated with Q_N and Q_P, the capacitance of the wire that connects the output of the inverter to other circuits, and the input capacitance of the logic gates the inverter drives, by a single equivalent capacitance C connected between the output node of the inverter and ground.
2. Analyze the resulting capacitively loaded inverter to determine its t_{PLH} and t_{PHL}, and hence t_P.

We shall study these two separable steps in reverse order. Thus, in this section we show how the propagation delay can be determined. Then, in Section 14.4.3, we show how to calculate the value of C.

Figure 14.30(a) shows a CMOS inverter with a capacitance C connected between its output node and ground. To determine the propagation delays t_{PHL} and t_{PLH}, we apply to the input an ideal pulse, that is, one with zero rise and fall times, as shown in Fig. 14.30(b). Since the circuit has a symmetric structure, the analyses to determine the two propagation delays will be similar. Therefore, we will derive t_{PHL} in detail and extrapolate the result to determine t_{PLH}.

Just prior to the leading edge of the input pulse (i.e., at $t = 0-$), the output voltage is equal to V_{DD} and capacitor C is charged to this voltage. At $t = 0$, v_I rises to V_{DD}, causing Q_P to turn off and Q_N to turn on. From then on, the circuit is equivalent to that shown in Fig. 14.30(c), with the initial value of $v_O = V_{DD}$. Thus, at $t = 0+$, Q_N will operate in the saturation region and will supply a relatively large current to begin the process of discharging C. Figure 14.30(d) shows the trajectory of the operating point of Q_N as C is discharged. Here we are interested in

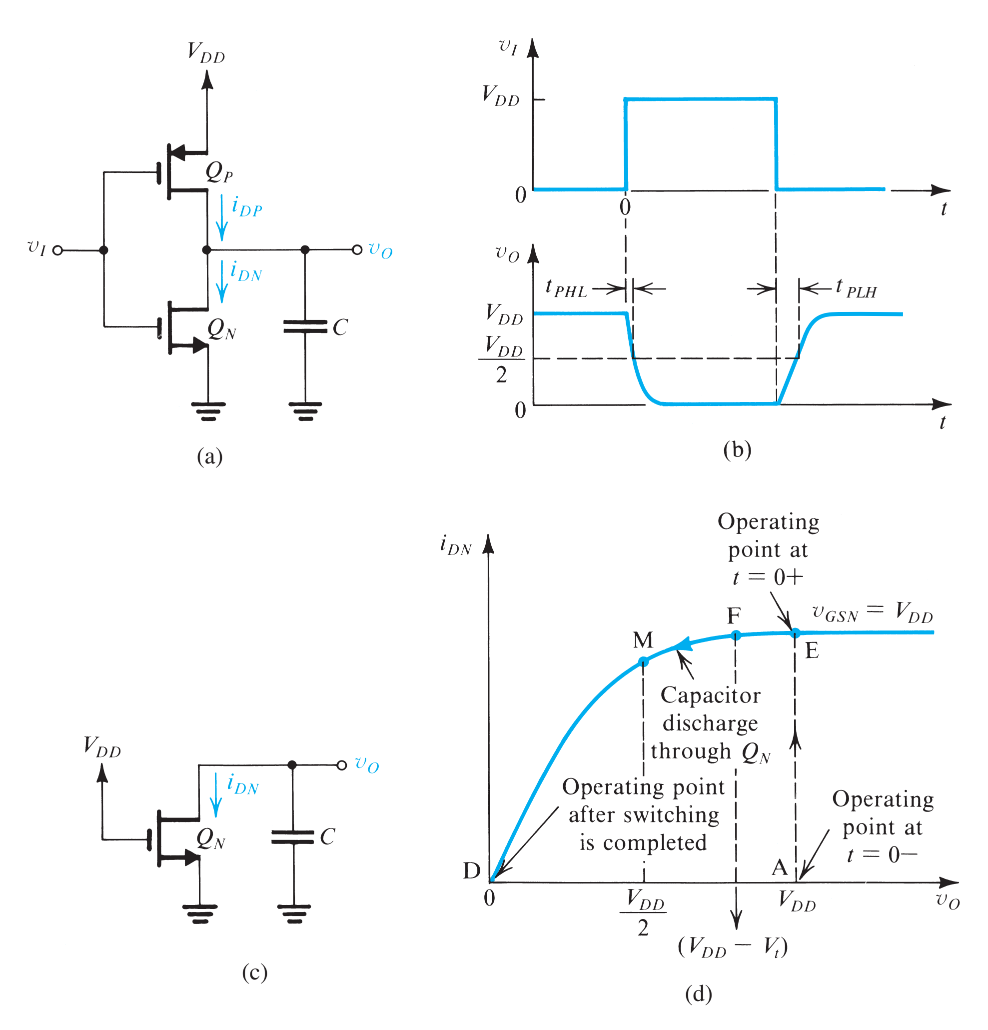


Figure 14.30 Dynamic operation of a capacitively loaded CMOS inverter: **(a)** circuit; **(b)** input and output waveforms; **(c)** equivalent circuit during the capacitor discharge; **(d)** trajectory of the operating point as the input goes high and C discharges through Q_N.

the interval t_{PHL} during which v_O reduces from V_{DD} to $V_{DD}/2$. Correspondingly, the operating point of Q_N moves from E to M. For a portion of this time, corresponding to the segment EF of the trajectory, Q_N operates in saturation. Then at F, $v_O = V_{DD} - V_t$, and Q_N enters the triode region.

A simple approach for determining t_{PHL} consists of first calculating the average value of the current supplied by Q_N over the segment EM. Then, we use this average value of the discharge current to determine t_{PHL} by means of the charge balance equation

$$I_{av} t_{PHL} = C[V_{DD} - (V_{DD}/2)]$$

resulting in

$$t_{PHL} = \frac{CV_{DD}}{2I_{av}} \tag{14.46}$$

The value of I_{av} can be found as follows:

$$I_{av} = \frac{1}{2}[i_{DN}(\text{E}) + i_{DN}(\text{M})] \tag{14.47}$$

where

$$i_{DN}(\text{E}) = \frac{1}{2}k'_n\left(\frac{W}{L}\right)_n(V_{DD} - V_{tn})^2 \tag{14.48}$$

and

$$i_{DN}(\text{M}) = k'_n\left(\frac{W}{L}\right)_n\left[(V_{DD} - V_{tn})\left(\frac{V_{DD}}{2}\right) - \frac{1}{2}\left(\frac{V_{DD}}{2}\right)^2\right] \tag{14.49}$$

Note that we have assumed $\lambda_n = 0$. Combining Eqs. (14.46) to (14.49) provides

$$t_{PHL} = \frac{\alpha_n C}{k'_n(W/L)_n V_{DD}} \tag{14.50}$$

where α_n is a factor determined by the relative values of V_t and V_{DD};

$$\alpha_n = 2\bigg/\left[\frac{7}{4} - \frac{3V_{tn}}{V_{DD}} + \left(\frac{V_{tn}}{V_{DD}}\right)^2\right] \tag{14.51}$$

The value of α_n typically falls in the range of 1 to 2.

An expression for the low-to-high inverter delay, t_{PLH}, can be written by analogy to the t_{PHL} expression in Eq. (14.50),

$$t_{PLH} = \frac{\alpha_p}{k'_p(W/L)_p V_{DD}} \tag{14.52}$$

where

$$\alpha_p = 2\bigg/\left[\frac{7}{4} - \frac{3|V_{tp}|}{V_{DD}} + \left|\frac{V_{tp}}{V_{DD}}\right|^2\right] \tag{14.53}$$

Finally, the propagation delay t_P can be found as the average of t_{PHL} and t_{PLH},

$$t_P = \frac{1}{2}(t_{PHL} + t_{PLH})$$

Examination of the formulas in Eqs. (14.50) to (14.53) enables us to make a number of useful observations:

1. As expected, the two components of t_P can be equalized by selecting the (W/L) ratios to equalize k_n and k_p, that is, by matching Q_N and Q_P. This assumes that $\alpha_n = \alpha_p$, which obtains when $V_{tn} = -V_{tp}$.

2. Since t_P is proportional to C, the designer should strive to reduce C. This is achieved by using the minimum possible channel length and by minimizing wiring and other parasitic capacitances. Careful layout of the chip can result in significant reduction in such capacitances.
3. Using a process technology with larger transconductance parameter k' can result in shorter propagation delays. Keep in mind, however, that for such processes C_{ox} is increased, and thus the value of C increases at the same time (more on this later).
4. Using larger W/L ratios can result in a reduction in t_P. Care, however, should be exercised here also, since increasing the size of the devices increases the value of C, and thus the expected reduction in t_P might not materialize. Reducing t_P by increasing W/L, however, is an effective strategy when C is dominated by components not directly related to the size of the driving device (such as wiring or fan-out devices).
5. A larger supply voltage V_{DD} results in a lower t_P. However, V_{DD} is determined by the process technology and thus is often not under the control of the designer. Furthermore, modern process technologies in which device sizes are reduced require lower V_{DD} (see Appendix K). A motivating factor for lowering V_{DD} is the need to keep the dynamic power dissipation at acceptable levels, especially in very-high-density chips. We will have more to say on this point in Section 14.6.

These observations clearly illustrate the conflicting requirements and the trade-offs available in the design of a CMOS digital integrated circuit (and indeed in any engineering design problem).

An Alternative Approach The formulas derived above for t_{PHL} and t_{PLH} underestimate the delay values for inverters implemented in deep-submicron technologies. This arises because of the velocity-saturation effect, which we shall discuss briefly in Section 15.1. There we will see that velocity saturation results in lower MOSFET currents in the saturation region, and hence in increased delay times. To deal with this problem, we present a very simple alternative approach to estimating the inverter propagation delay.

Figure 14.31 illustrates the alternative approach. During the discharge delay t_{PHL}, Q_N is replaced by an equivalent resistance R_N. Similarly, during the charging delay t_{PLH}, Q_P is replaced by an equivalent resistance R_P. It is easy to show that

$$t_{PHL} = 0.69 R_N C \tag{14.54}$$

and

$$t_{PLH} = 0.69 R_P C \tag{14.55}$$

Empirical values have been found for R_N and R_P,

$$R_N = \frac{12.5}{(W/L)_n}\ \text{k}\Omega \tag{14.56}$$

$$R_P = \frac{30}{(W/L)_p}\ \text{k}\Omega \tag{14.57}$$

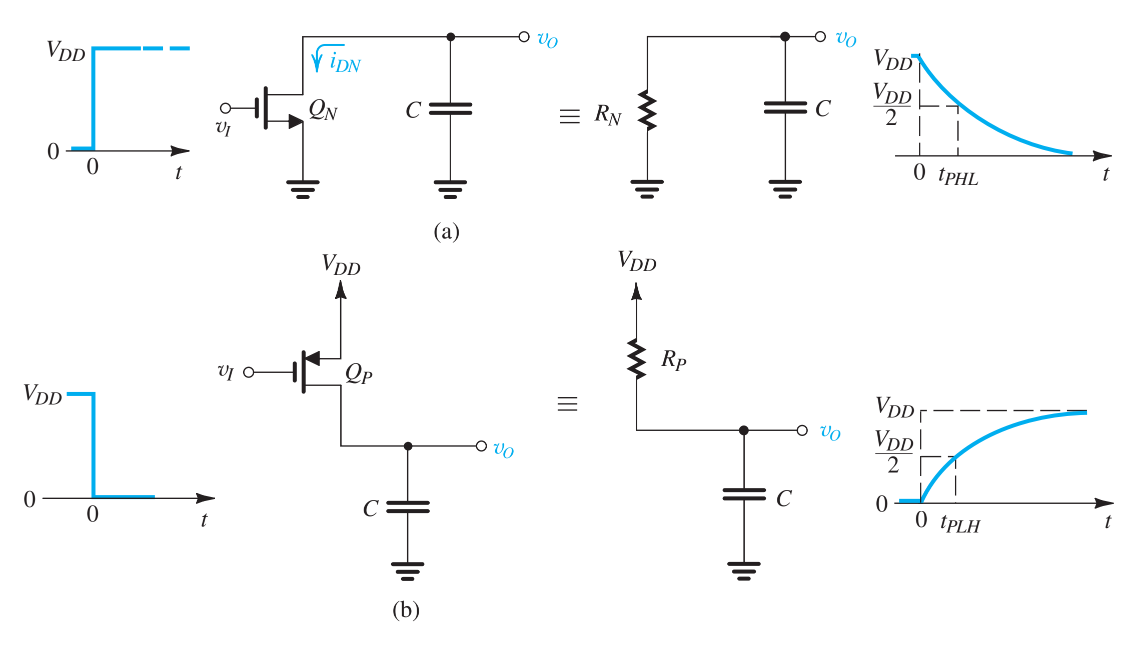


Figure 14.31 Equivalent circuits for determining the propagation delays **(a)** t_{PHL} and **(b)** t_{PLH} of the inverter.

Furthermore, it has been found that these values apply for a number of CMOS fabrication processes including 0.25 μm, 0.18 μm, and 0.13 μm (see Hodges et al., 2004).

As a final point, we note that the delay expressions in Eqs. (14.54) and (14.55) are obtained by assuming that the inverter is driven by a step-input voltage. In the more practical case of a ramp-input voltage, it has been shown that the 0.69 factor approaches unity, thus

$$t_{PHL} \simeq R_n C \qquad (14.54')$$

and

$$t_{PLH} \simeq R_p C \qquad (14.55')$$

Example 14.6 Determining the Propagation Delay of the CMOS Inverter

For the 0.25-μm process characterized by $V_{DD} = 2.5$ V, $V_{tn} = -V_{tp} = 0.5$ V, $k_n' = 3.5k_p' = 110\ \mu\text{A/V}^2$, find t_{PLH}, t_{PHL}, and t_P for an inverter for which $(W/L)_n = 1.5$ and $(W/L)_p = 3$, and for $C = 10$ fF. Use both the approach based on average currents and that based on equivalent resistances, and compare the results obtained. If to save on power dissipation, the inverter is operated at $V_{DD} = 2.0$ V, by what factor does t_P change?

Example 14.6 *continued*

Solution

(a) Using the average current approach, we determine from Eq. (14.51),

$$\alpha_n = \frac{2}{\frac{7}{4} - \frac{3 \times 0.5}{2.5} + \left(\frac{0.5}{2.5}\right)^2} = 1.7$$

and using Eq. (14.50),

$$t_{PHL} = \frac{1.7 \times 10 \times 10^{-15}}{110 \times 10^{-6} \times 1.5 \times 2.5} = 41.2 \text{ ps}$$

Since $|V_{tp}| = V_{tn}$,

$$\alpha_p = \alpha_n = 1.7$$

and we can determine t_{PLH} from Eq. (14.52) as

$$t_{PLH} = \frac{1.7 \times 10 \times 10^{-15}}{(110/3.5) \times 10^{-6} \times 3 \times 2.5} = 72.1 \text{ ps}$$

The propagation delay can now be found as

$$t_P = \frac{1}{2}(t_{PHL} + t_{PLH})$$

$$= \frac{1}{2}(41.2 + 72.1) = 56.7 \text{ ps}$$

(b) Using the equivalent-resistance approach, we first find R_N from Eq. (14.56) as

$$R_N = \frac{12.5}{1.5} = 8.33 \text{ k}\Omega$$

and then use Eq. (14.54) to determine t_{PHL},

$$t_{PHL} = 0.69 \times 8.33 \times 10^3 \times 10 \times 10^{-15} = 57.5 \text{ ps}$$

Similarly we use Eq. (14.57) to determine R_P,

$$R_P = \frac{30}{3} = 10 \text{ k}\Omega$$

and Eq. (14.55) to determine t_{PLH},

$$t_{PLH} = 0.69 \times 10 \times 10^3 \times 10 \times 10^{-15} = 69 \text{ ps}$$

Thus, while the value obtained for t_{PHL} is higher than that found using average currents, the value for t_{PLH} is about the same. Finally, t_P can be found as

$$t_P = \frac{1}{2}(57.5 + 69) = 63.2 \text{ ps}$$

which is a little higher than the value found using average currents.

To find the change in propagation delays obtained when the inverter is operated at $V_{DD} = 2.0$ V, we have to use the method of average currents. (The dependence on the power-supply voltage is absorbed in the empirical values of R_N and R_P.) Using Eq. (14.51), we write

$$\alpha_n = \frac{2}{\frac{7}{4} - \frac{3 \times 0.5}{2} + \left(\frac{0.5}{2}\right)^2} = 1.9$$

The value of t_{PHL} can now be found by using Eq. (14.50):

$$t_{PHL} = \frac{1.9 \times 10 \times 10^{-15}}{110 \times 10^{-6} \times 1.5 \times 2} = 57.6 \text{ ps}$$

Similarly, the value of $\alpha_p = \alpha_n = 1.9$ can be substituted in Eq. (14.52) to obtain,

$$t_{PLH} = \frac{1.9 \times 10 \times 10^{-15}}{(110/3.5) \times 10^{-6} \times 3 \times 2} = 100.8 \text{ ps}$$

and t_P can be calculated as

$$t_P = \frac{1}{2}(57.6 + 100.8) = 79.8 \text{ ps}$$

Thus, as expected, reducing V_{DD} has resulted in increased propagation delay.

Before leaving the subject of propagation delay, we should emphasize that hand analysis using the simple formulas above should *not* be expected to yield precise results. Rather, its value is in obtaining design insight. Precise results can always be obtained using SPICE and Multisim simulations (see examples in Appendix B and the extensive material on the website). However, it is never a good idea to use simulation if one does not know beforehand approximate values of the expected results.

EXERCISES

14.11 For a CMOS inverter fabricated in a 0.18-μm process with $V_{DD} = 1.8$ V, $V_{tn} = -V_{tp} = 0.5$ V, $k'_n = 4k'_p = 300\ \mu\text{A/V}^2$, and having $(W/L)_n = 1.5$ and $(W/L)_p = 3$, find t_{PHL}, t_{PLH}, and t_P when the equivalent load capacitance $C = 10$ fF. Use the method of average currents.
Ans. 24.8 ps; 49.6 ps; 37.2 ps

D14.12 For a CMOS inverter fabricated in a 0.13-μm process, use the equivalent-resistances approach to determine $(W/L)_n$ and $(W/L)_p$ so that $t_{PLH} = t_{PHL} = 50$ ps when the effective load capacitance $C =$ 20 fF.
Ans. 3.5; 8.3

14.4.3 Determining the Equivalent Load Capacitance C

Having determined the propagation delay of the CMOS inverter in terms of the equivalent load capacitance C, it now remains to determine the value of C. For this purpose, a thorough understanding of the various capacitances in a MOS transistor is essential, and we urge the reader to review the material in Section 10.2.1.

Figure 14.32 shows the circuit for determining the propagation delay of the CMOS inverter formed by Q_1 and Q_2. Note that we are showing the inverter driving a similar inverter formed by transistors Q_3 and Q_4. This reflects a practical situation and will help us explain how to determine the contribution of a driven inverter to the equivalent capacitance C at the output of the inverter under study (that formed by Q_1 and Q_2).

Indicated in Fig. 14.32 are the various transistor capacitances that connect to the output node of the Q_1–Q_2 inverter. Also shown is the **wiring capacitance** C_w, which represents

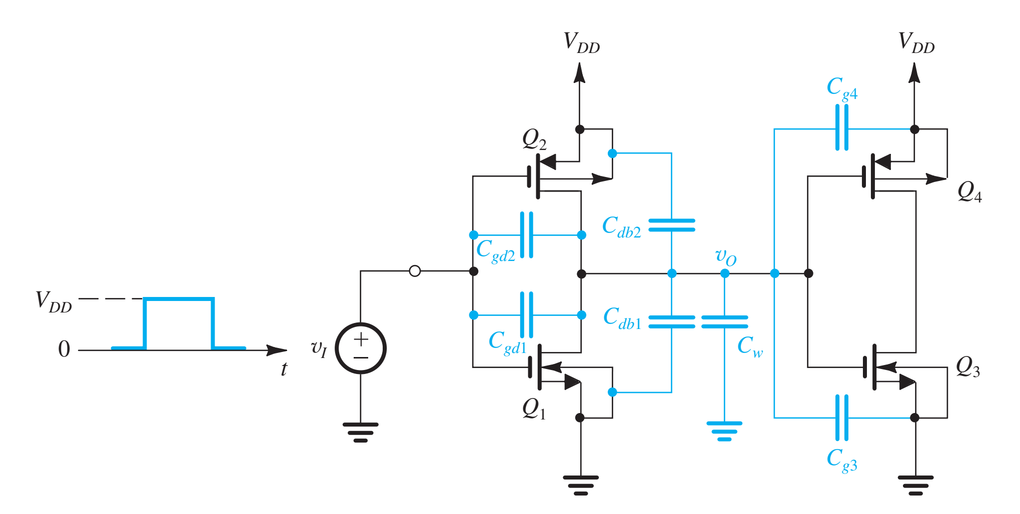


Figure 14.32 Circuit for analyzing the propagation delay of the inverter formed by Q_1 and Q_2, which is driving a similar inverter formed by Q_3 and Q_4.

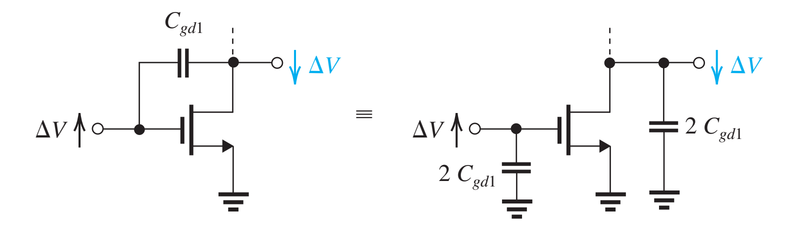


Figure 14.33 The Miller multiplication of the feedback capacitance C_{gd1}.

the capacitance of the wire or **interconnect** that connects the output of the Q_1-Q_2 inverter to the input of the Q_3-Q_4 inverter. Interconnect capacitances have become increasingly dominant as the technology has scaled down. In fact, some digital IC designers hold the view that interconnect poses a greater limitation on the speed of operation than the transistors themselves. We will discuss this topic briefly in Section 15.1.

A glance at the circuit in Fig. 14.32 should be sufficient to indicate that a pencil-and-paper analysis is virtually impossible. That, of course, is the reason we opted for the simplification of replacing all these capacitances with an equivalent capacitance C. Before we consider the determination of C, it is useful to observe that during t_{PLH} or t_{PHL}, the output of the first inverter changes from 0 to $V_{DD}/2$ or from V_{DD} to $V_{DD}/2$, respectively. Assuming that the switching threshold of the second inverter is $V_{DD}/2$, it follows that the second inverter remains in the same state during each of our analysis intervals. This observation will have an important bearing on our estimation of the equivalent input capacitance of the second inverter. Let's now consider the contribution of each of the capacitances in Fig. 14.32 to the value of the equivalent load capacitance C:

1. The gate–drain overlap capacitance of Q_1, C_{gd1}, can be replaced by an equivalent capacitance between the output node and ground of $2C_{gd1}$. The factor 2 arises because of the Miller effect (Section 10.3.3). Specifically, refer to Fig. 14.33 and note that as v_I goes high and v_O goes low by the same amount, the change in voltage across C_{gd1} is twice that amount. Thus the output node sees in effect twice the value of C_{gd1}. The same applies for the gate–drain overlap capacitance of Q_2, C_{gd2}, which can be replaced by a capacitance $2C_{gd2}$ between the output node and ground.
2. Each of the drain–body capacitances C_{db1} and C_{db2} has a terminal at a constant voltage. Thus for the purpose of our analysis here, C_{db1} and C_{db2} can be replaced with equal capacitances between the output node and ground. Note, however, that the formulas given in Section 10.2.1 for calculating C_{db1} and C_{db2} are small-signal relationships, whereas the analysis here is obviously a large-signal one. A technique has been developed for finding equivalent large-signal values for C_{db1} and C_{db2} (see Hodges et al., 2004 and Rabaey et al., 2003).
3. Since the second inverter does not switch states, we will assume that the input capacitances of Q_3 and Q_4 remain approximately constant and equal to the total gate capacitance $\left(WLC_{ox}+C_{gsov}+C_{gdov}\right)$. That is, the input capacitance of the load inverter will be

$$C_{g3}+C_{g4}=(WL)_3C_{ox}+(WL)_4C_{ox}+C_{gsov3}+C_{gdov3}+C_{gsov4}+C_{gdov4} \qquad (14.58)$$

4. The last component of C is the wiring capacitance C_w, which simply adds to the value of C.

Thus, the total value of C is given by

$$C = 2C_{gd1} + 2C_{gd2} + C_{db1} + C_{db2} + C_{g3} + C_{g4} + C_w \tag{14.59}$$

Example 14.7 Determining the Effective Load Capacitance C and the Propagation Delay

Consider a CMOS inverter fabricated in a 0.25-μm process for which $C_{ox} = 6$ fF/μm^2, $\mu_n C_{ox} = 110$ μA/V^2, $\mu_p C_{ox} = 30$ μA/V^2, $V_{tn} = -V_{tp} = 0.5$ V, and $V_{DD} = 2.5$ V. The W/L ratio of Q_N is 0.375 μm/0.25 μm, and that for Q_P is 1.125 μm/0.25 μm. The gate–source and gate–drain overlap capacitances are specified to be 0.3 fF/μm of gate width. Further, the effective (large-signal) values of drain–body capacitances are $C_{dbn} = 1$ fF and $C_{dbp} = 1$ f F. The wiring capacitance $C_w = 0.2$ f F. Find t_{PHL}, t_{PLH}, and t_P when the inverter is driving an identical inverter.

Solution

First, we determine the value of the equivalent capacitance C using Eqs. (14.58) and (14.59),

$$C = 2C_{gd1} + 2C_{gd2} + C_{db1} + C_{db2} + C_{g3} + C_{g4} + C_w$$

where

$$C_{gd1} = 0.3 \times W_n = 0.3 \times 0.375 = 0.1125 \text{ fF}$$
$$C_{gd2} = 0.3 \times W_p = 0.3 \times 1.125 = 0.3375 \text{ fF}$$
$$C_{db1} = 1 \text{ fF}$$
$$C_{db2} = 1 \text{ fF}$$
$$C_{g3} = 0.375 \times 0.25 \times 6 + 2 \times 0.3 \times 0.375 = 0.7875 \text{ fF}$$
$$C_{g4} = 1.125 \times 0.25 \times 6 + 2 \times 0.3 \times 1.125 = 2.3625 \text{ fF}$$
$$C_w = 0.2 \text{ fF}$$

Thus,

$$C = 2 \times 0.1125 + 2 \times 0.3375 + 1 + 1 + 0.7875 + 2.3625 + 0.2 = 6.25 \text{ fF}$$

Next we use Eqs. (14.51) and (14.52) to determine t_{PHL},

$$\alpha_n = \frac{2}{\frac{7}{4} - \frac{3 \times 0.5}{2.5} + \left(\frac{0.5}{2.5}\right)^2} = 1.7$$

$$t_{PHL} = \frac{1.7 \times 6.25 \times 10^{-15}}{110 \times 10^{-6} \times (0.375/0.25) \times 2.5} = 25.8 \text{ ps}$$

Similarly, we use Eqs. (14.53) and (14.54) to determine t_{PLH},

$$\alpha_p = 1.7$$

$$t_{PLH} = \frac{1.7 \times 6.25 \times 10^{-15}}{30 \times 10^{-6} \times (1.125/0.25) \times 2.5} = 31.5 \text{ ps}$$

Finally, we determine t_P as

$$t_P = \frac{1}{2}(25.8 + 31.5) = 28.7 \text{ ps}$$

EXERCISES

14.13 Consider the inverter specified in Example 14.7 when loaded with an additional 0.1-pF capacitance. What will the propagation delay become?
Ans. 488 ps

14.14 In an attempt to decrease the area of the inverter in Example 14.7, $(W/L)_p$ is made equal to $(W/L)_n$. What is the percentage reduction in area achieved? Find the new values of C, t_{PHL}, t_{PLH}, and t_P. Assume that C_{dbp} does not change significantly.
Ans. 50%; 4.225 fF; 17.4 ps; 63.8 ps; 40.6 ps

14.15 For the inverter of Example 14.7, find the theoretical maximum frequency at which it can be operated.
Ans. 17.4 GHz

14.5 Transistor Sizing

In this section we address the extremely important design question of selecting appropriate sizes (i.e., L and W/L values) for all transistors in a CMOS logic circuit. We begin with the CMOS inverter and then consider general logic gates.

14.5.1 Inverter Sizing

In this section we are concerned with the selection of appropriate values for the channel length L and the (W/L) ratios for the two transistors Q_N and Q_P in an inverter. Our reasoning can be summarized as follows.

1. To minimize area, the length of all channels is usually made equal to the minimum length permitted by the given technology.

2. In a given inverter, if our interest is strictly to minimize area, $(W/L)_n$ is usually selected in the range 1 to 1.5. The selection of $(W/L)_p$ relative to $(W/L)_n$ has influence on the noise margins and t_{PLH}. Both are optimized by matching Q_P and Q_N. This, however, is usually wasteful of area and, equally important, can increase the effective capacitance C, so that although t_{PLH} is made equal to t_{PHL}, the value of both can be higher than in the case without matching (see Problem 14.55). Thus, selecting $(W/L)_p = (W/L)_n$ is a possibility, and $(W/L)_p = 2(W/L)_n$ is a frequently used compromise.
3. Having settled on an appropriate ratio of $(W/L)_p$ to $(W/L)_n$, we still have to select $(W/L)_n$ to reduce t_P and thus allow higher speeds of operation. Any increase in $(W/L)_n$ and proportionally in $(W/L)_p$ will of course increase area, and hence the inverter contribution to the value of the equivalent capacitance C. To be more precise we express C as the sum of an intrinsic component C_{int} contributed by Q_N and Q_P of the inverter, and an extrinsic component C_{ext} resulting from the wiring and the input capacitance of the driven gates,

$$C = C_{\text{int}} + C_{\text{ext}} \tag{14.60}$$

Increasing $(W/L)_n$ and $(W/L)_p$ of the inverter by a factor S relative to that of a minimum-size inverter for which $C_{\text{int}} = C_{\text{int0}}$ results in

$$C = SC_{\text{int0}} + C_{\text{ext}} \tag{14.61}$$

Now, if we use the equivalent-resistances approach to compute t_P and define an equivalent inverter resistance R_{eq} as

$$R_{\text{eq}} = \frac{1}{2}(R_N + R_P) \tag{14.62}$$

then,

$$t_P = 0.69 R_{\text{eq}} C \tag{14.63}$$

Further, if for the minimum-size inverter R_{eq} is R_{eq0}, increasing $(W/L)_n$ and $(W/L)_p$ by the factor S reduces R_{eq} by the same factor:

$$R_{\text{eq}} = R_{\text{eq0}}/S \tag{14.64}$$

Combining Eqs. (14.63), (14.64), and (14.61), we obtain

$$t_P = 0.69\left(\frac{R_{\text{eq0}}}{S}\right)(SC_{\text{int0}} + C_{\text{ext}})$$

$$t_P = 0.69\left(R_{\text{eq0}} C_{\text{int0}} + \frac{1}{S} R_{\text{eq0}} C_{\text{ext}}\right) \tag{14.65}$$

We thus see that scaling the W/L ratios does *not* change the component of t_P caused by the capacitances of Q_N and Q_P. It does, however, reduce the component of t_P that results from capacitances external to the inverter itself. It follows that one can use Eq. (14.65) to decide on a suitable scaling factor S that keeps t_P below a specified maximum value, keeping in mind of course the effect of increasing S on silicon area.

EXERCISE

14.16 For the inverter analyzed in Example 14.7:
(a) Find the intrinsic and extrinsic components of C.
(b) By what factor must $(W/L)_n$ and $(W/L)_p$ be increased to reduce the extrinsic part of t_P by a factor of 2?
(c) Estimate the resulting t_P.
(d) By what factor is the inverter area increased?
Ans. (a) 2.9 fF, 3.35 fF; (b) 2; (c) 21 ps; (d) 2

FEDERICO FAGGIN—A PIONEER IN MICROPROCESSOR ELECTRONICS:

Holder of a degree in physics from the University of Padua, Federico Faggin first worked for SGS-Fairchild in Italy. In 1968 he relocated to California, joining Fairchild in Palo Alto, where he developed the silicon-gate MOS device that has dominated MOS production ever since. In 1970 he joined Intel, where he led the design and production of the Intel 4004, the world's first commercial single-chip microcomputer, introduced in 1971. This design was based on a four-chip version with separated memory that Ted Hoff had designed in 1969 in response to the request by a Japanese calculator company for a twelve-chip flexible design. The 4004, a 4-bit processor, included 2300 PMOS logic circuits (a long-obsolete logic-circuit form) on a 3 mm × 4 mm die, using a random-logic process created by Faggin. Subsequently, Faggin was responsible for the 8008 at Intel, and the Z80 8-bit microprocessor at Zilog (which he founded in 1974). Later, he went on to cofound several other companies, including Synaptics (in 1986 with Carver Mead and others), which provided touch-sensitive pad and screen designs for the mobile and PC products of many manufacturers.

14.5.2 Transistor Sizing in CMOS Logic Gates

Once a CMOS gate circuit has been generated, the only significant step remaining in the design is to decide on W/L ratios for all devices. These ratios usually are selected to provide the gate with current-driving capability in both directions equal to that of the basic inverter. For the basic inverter design, denote $(W/L)_n = n$ and $(W/L)_p = p$, where n is usually 1 to 1.5 and, for a matched design, $p = (\mu_n/\mu_p)\, n$; it should be noted, however, that often $p = 2n$ and for minimum area $p = n$. Thus, we wish to select individual W/L ratios for all transistors in a logic gate so that the PDN should be able to provide a capacitor discharge current *at least* equal to that of an NMOS transistor with $W/L = n$, and the PUN should be able to provide a charging current *at least* equal to that of a PMOS transistor with $W/L = p$. This will guarantee a *worst-case* gate delay equal to that of the basic inverter.[4]

[4]This statement assumes that the total effective capacitance C of the logic gate is the same as that of the inverter. In actual practice, the value of C will be larger for a gate, especially as the fan-in is increased.

In the preceding description, the idea of "worst case" should be emphasized. It means that in deciding on device sizing, we should find the input combinations that result in the lowest output current and then choose sizes that will make this current equal to that of the basic inverter. Before we consider examples, we need to address the issue of determining the current-driving capability of a circuit consisting of a number of MOS devices. In other words, we need to find the *equivalent W/L ratio* of a network of MOS transistors. Toward that end, we consider the parallel and series connection of MOSFETs and find the equivalent W/L ratios.

The derivation of the equivalent W/L ratio is based on the fact that the on-resistance of a MOSFET is inversely proportional to W/L (see Eqs. 14.56 and 14.57). Thus, if a number of MOSFETs having ratios of $(W/L)_1$, $(W/L)_2$, . . ., are connected in series, the equivalent series resistance obtained by adding the on-resistances will be

$$
\begin{aligned}
R_{\text{series}} &= R_{N1} + R_{N2} + \dots \\
&= \frac{\text{constant}}{(W/L)_1} + \frac{\text{constant}}{(W/L)_2} + \dots \\
&= \text{constant}\left[\frac{1}{(W/L)_1} + \frac{1}{(W/L)_2} + \dots\right] \\
&= \frac{\text{constant}}{(W/L)_{\text{eq}}}
\end{aligned}
$$

resulting in the following expression for $(W/L)_{\text{eq}}$ for transistors connected in series:

$$(W/L)_{\text{eq}} = \frac{1}{\dfrac{1}{(W/L)_1} + \dfrac{1}{(W/L)_2} + \dots} \tag{14.66}$$

Similarly, we can show that the parallel connection of transistors with W/L ratios of $(W/L)_1$, $(W/L)_2$,. . ., results in an equivalent W/L of

$$(W/L)_{\text{eq}} = (W/L)_1 + (W/L)_2 + \dots \tag{14.67}$$

As an example, two identical MOS transistors with individual W/L ratios of 4 result in an equivalent W/L of 2 when connected in series and of 8 when connected in parallel.[5]

As an example of proper sizing, consider the four-input NOR in Fig. 14.34. Here, the worst case (the lowest current) for the PDN is obtained when only one of the NMOS transistors is conducting. We therefore select the W/L of each NMOS transistor to be equal to that of the NMOS transistor of the basic inverter, namely, n. For the PUN, however, the worst-case situation (and indeed the only case) occurs when all inputs are low and the four series PMOS transistors are conducting. Since the equivalent W/L will be one-quarter of that of each PMOS device, we should select the W/L ratio of each PMOS transistor to be four times that of Q_P of the basic inverter, that is, $4p$.

[5]Another way of thinking about this is as follows: Connecting MOS transistors in series is equivalent to adding the lengths of their channels while the width does not change; connecting MOS transistors in parallel does not change the channel length but increases the width to the sum of the W's.

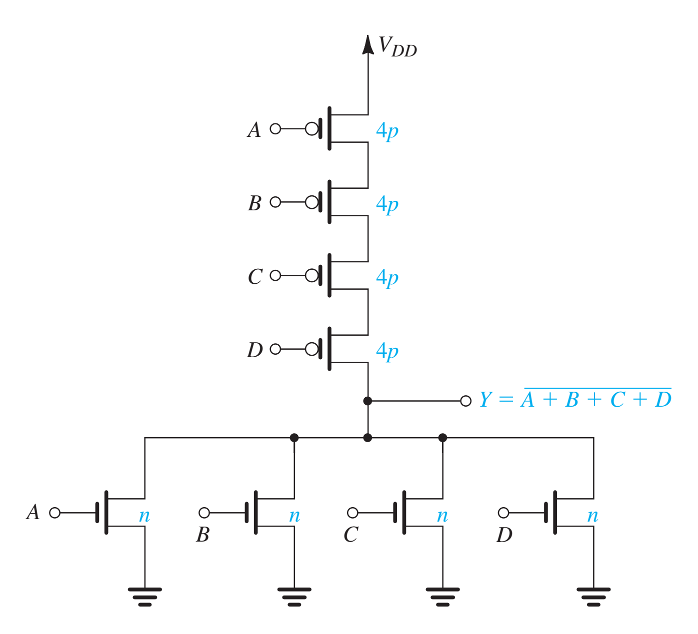


Figure 14.34 Proper transistor sizing for a four-input NOR gate. Note that *n* and *p* denote the *W/L* ratios of Q_N and Q_P, respectively, of the basic inverter.

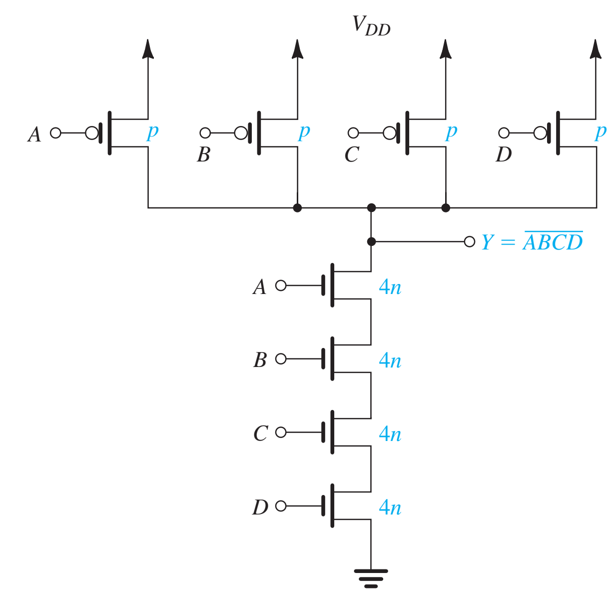


Figure 14.35 Proper transistor sizing for a four-input NAND gate. Note that *n* and *p* denote the *W/L* ratios of Q_N and Q_P, respectively, of the basic inverter.

As another example, we show in Fig. 14.35 the proper sizing for a four-input NAND gate. Comparison of the NAND and NOR gates in Figs. 14.34 and 14.35 indicates that because *p* is usually two to three times *n*, the NOR gate will require much greater area than the NAND gate. For this reason, NAND gates are generally preferred for implementing combinational-logic functions in CMOS.

Example 14.8 Transistor Sizing of a CMOS Gate

Provide transistor W/L ratios for the logic circuit shown in Fig. 14.36. Assume that for the basic inverter $n = 1.5$ and $p = 5$ and that the channel length is 0.25 μm.

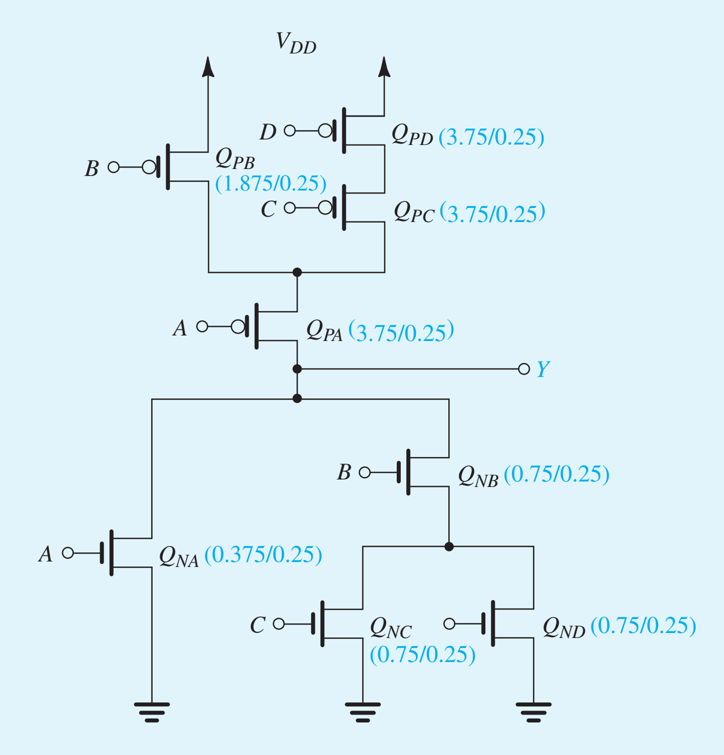


Figure 14.36 Circuit for Example 14.8.

Solution

Refer to Fig. 14.36, and consider the PDN first. We note that the worst case occurs when Q_{NB} is on and either Q_{NC} or Q_{ND} is on. That is, in the worst case, we have two transistors in series. Therefore, we select each of Q_{NB}, Q_{NC}, and Q_{ND} to have twice the width of the n-channel device in the basic inverter, thus

$$Q_{NB}\text{: } W/L = 2n = 3 = 0.75/0.25$$

$$Q_{NC}\text{: } W/L = 2n = 3 = 0.75/0.25$$

$$Q_{ND}\text{: } W/L = 2n = 3 = 0.75/0.25$$

For transistor Q_{NA}, select W/L to be equal to that of the n-channel device in the basic inverter:

$$Q_{NA}: W/L = n = 1.5 = 0.375/0.25$$

Next, consider the PUN. Here, we see that in the worst case, we have three transistors in series: Q_{PA}, Q_{PC}, and Q_{PD}. Therefore, we select the W/L ratio of each of these to be three times that of Q_P in the basic inverter, that is, $3p$, thus

$$Q_{PA}: W/L = 3p = 15 = 3.75/0.25$$

$$Q_{PC}: W/L = 3p = 15 = 3.75/0.25$$

$$Q_{PD}: W/L = 3p = 15 = 3.75/0.25$$

Finally, the W/L ratio for Q_{PB} should be selected so that the equivalent W/L of the series connection of Q_{PB} and Q_{PA} should be equal to p. It follows that for Q_{PB} the ratio should be $1.5p$,

$$Q_{PB}: W/L = 1.5p = 7.5 = 1.875/0.25$$

Figure 14.36 shows the circuit with the transistor sizes indicated.

14.5.3 Effects of Fan-In and Fan-Out on Propagation Delay

Each additional input to a CMOS gate requires two additional transistors, one NMOS and one PMOS. This is in contrast to other forms of MOS logic, where each additional input requires only one additional transistor, such as in the pseudo-NMOS logic, whose basic inverter was considered in Example 14.3 and which will be studied in some detail in the next chapter. The additional transistor in CMOS not only increases the chip area but also increases the total effective capacitance per gate and in turn increases the propagation delay. The size-scaling method described earlier compensates for some (but not all) of the increase in t_P. Specifically, by increasing device size, we are able to preserve the current-driving capability. However, the capacitance C increases because of both the increased number of inputs and the increase in device size. Thus t_P will still increase with fan-in, a fact that imposes a practical limit on the fan-in of, say, the NAND gate to about 4. If a higher number of inputs is required, then "clever" logic design should be adopted to realize the given Boolean function with gates of no more than four inputs. This would usually mean an increase in the number of cascaded stages and thus an increase in delay. However, such an increase in delay can be less than the increase due to the large fan-in (see Problem 14.59).

An increase in a gate's fan-out adds directly to its load capacitance and, thus, increases its propagation delay.

Thus although CMOS has many advantages, it does suffer from increased circuit complexity when the fan-in and fan-out are increased, and from the corresponding effects of this complexity on both chip area and propagation delay. In Chapter 15, we shall study some simplified forms of CMOS logic that attempt to reduce this complexity, although at the expense of forgoing some of the advantages of basic CMOS.

EXERCISES

14.17 For a process technology with $L = 0.18\ \mu\text{m}$, $n = 1.5$, $p = 3$, give the sizes of all transistors in (a) a four-input NOR and (b) a four-input NAND. Also, give the relative areas of the two gates.

Ans.
(a) NMOS devices: $W/L = 0.27/0.18$, PMOS devices: 2.16/0.18;
(b) NMOS devices: $W/L = 1.08/0.18$, PMOS devices: 0.54/0.18;
NOR area/NAND area = 1.5

14.18 For the scaled NAND gate in Exercise 14.17, find the ratio of the maximum to minimum current available to (a) charge a load capacitance and (b) discharge a load capacitance.

Ans. (a) 4; (b) 1

14.5.4 Driving a Large Capacitance

In many cases in digital CMOS design, a logic gate must drive a large load capacitance. This might, for example, be due to a long wire on a chip, or to a requirement to drive an off-chip printed-circuit board trace, where the load capacitance can be several hundred times larger than the parasitic capacitances of the driving gate.

Let's investigate how to drive such a large load capacitance without causing the propagation delay to be unacceptably large. Figure 14.37(a) shows the large capacitive load C_L driven by a standard inverter. Note that we have simplified the model of the inverter by assuming that all its capacitances can be lumped into a capacitance C between its input and ground and that it has an effective output resistance R. Connecting C_L directly to the inverter output results in a propagation delay, assuming a ramp input, equal to the time constant τ,

$$t_P = \tau = C_L R \tag{14.68}$$

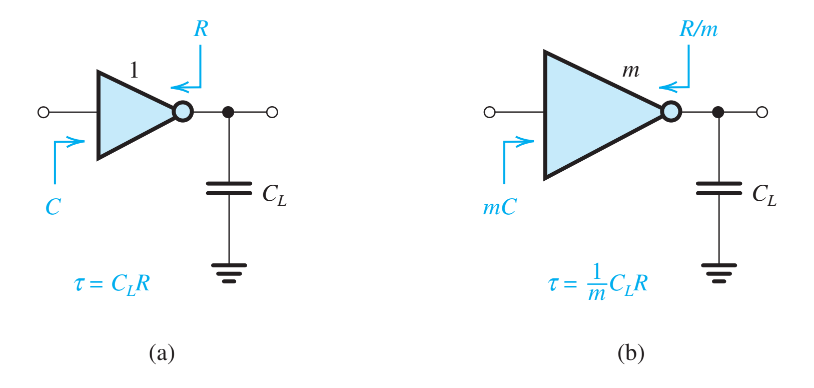
This propagation delay can be very large.

In an attempt to reduce the propagation delay, we can make the driver inverter large. Such a case is shown in Fig. 14.37(b), where an inverter m times larger than the standard inverter is used. Its output resistance will be R/m, that is, m times lower than that of the standard inverter. As a result, the propagation delay in this case will be

$$\tau = C_L(R/m) = \frac{1}{m}C_L R \tag{14.69}$$

which as desired has been reduced by a factor m. However, all is not well. Observe that the input capacitance of the large inverter is mC, which can be very large, requiring a large driving inverter to ensure that it does not contribute significantly to lengthening the overall propagation delay. Thus, it appears that we have not solved the problem, but rather shifted the burden to another inverter to drive the input of our large inverter.

The above reasoning leads to the idea of a chain of inverters connected in cascade, as shown in Fig. 14.37(c). Here we have n inverters of progressively larger sizes. In fact, it has been found that the optimum (i.e., lowest overall propagation delay) is obtained when each inverter in the chain is larger than the preceding inverter by the same factor x. Thus if inverter 1 has a unit



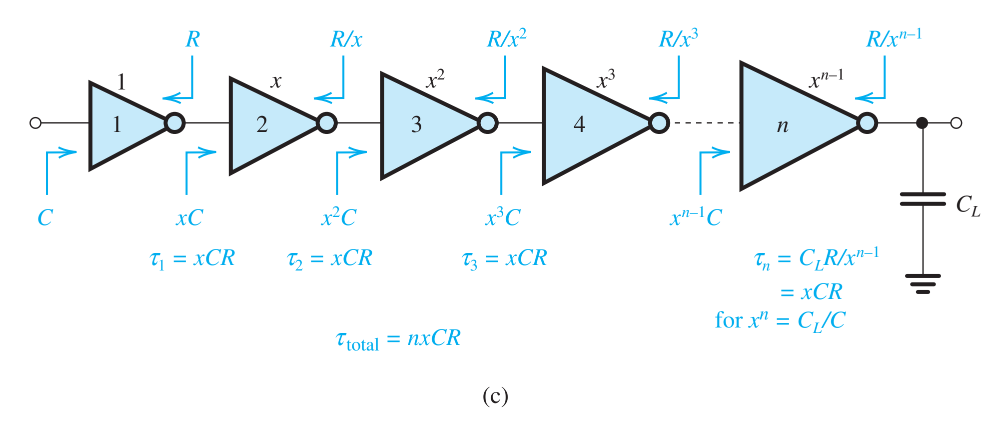


Figure 14.37 Driving a large load capacitance C_L: **(a)** directly; **(b)** by utilizing a large inverter; **(c)** by using a chain of progressively larger inverters.

size, inverter 2 has a size x, inverter 3 has a size x^2, and so on. Figure 14.37(c) shows the effect of inverter size scaling on its input capacitance and its equivalent output resistance. Observe that the delay time associated with the interface between each two succeeding inverters is $\tau = xCR$; that is, each interface contributes equally to the overall delay. This, of course, is a result of the geometric size scaling of the inverters in this chain. It has been shown that minimum delay is obtained if this equality of time constants extends to the output node, that is, by making

$$\tau_n \equiv C_L\left(\frac{R}{x^{n-1}}\right)$$

equal to xCR, which can be achieved if

$$x^n = \frac{C_L}{C} \tag{14.70}$$

in which case the overall delay becomes

$$t_P = \tau_{\text{total}} = nxCR \tag{14.71}$$

The question of selecting values for x and n remains. First, observe that there is already one condition on their values, namely, that in Eq. (14.70). It can be shown mathematically that the second condition that leads to minimum propagation delay (see Problem 14.62) is

$$x = e = 2.718 \tag{14.72}$$

In practice, it has been found that values for x between 2.5 and 4 lead to optimum performance (see Hodges et al., 2004).

Example 14.9 Design of an Inverter Chain to Drive a Large Load Capacitance

An inverter whose input capacitance $C = 10$ fF and whose equivalent output resistance $R = 1$ kΩ must ultimately drive a load capacitance $C_L = 1$ pF.

(a) What is the time delay that results if the inverter is connected directly to C_L?
(b) If a driver chain such as that in Fig. 14.37(c) is used, how many inverters n and what size ratio x should you use to minimize the total delay? What is the total path delay achieved?

Solution

(a) $t_P = \tau = C_L R = 10^{-12} \times 10^3 = 1$ ns.
(b) The delay is minimized by selecting

$$x = e = 2.718$$

and

$$x^n = \frac{C_L}{C} = \frac{10^{-12}}{10 \times 10^{-15}} = 100$$

which yields

$$n = \frac{\ln 100}{\ln x} = \frac{\ln 100}{\ln e} = 4.6$$

Since we must use an integral number of inverters, we select

$$n = 5$$

and obtain x from

$$x^n = x^5 = \frac{C_L}{C} = 100$$

which yields

$$x = (100)^{1/5} = 2.51$$

The total path delay will be

$$t_P = nxCR$$
$$= 5 \times 2.51 \times 10 \times 10^{-15} \times 10^3 = 125.5 \text{ ps}$$

which is a reduction in delay by a factor of about 8!

14.6 Power Dissipation

Many of today's integrated circuits are battery powered. Some even rely on "scavenged" energy, therefore severely limiting the supply of power. Other high-performance circuits, such as those found at computer server farms, have heat-dissipation limitations. Also, the desire to pack an ever-increasing number of gates on an IC chip (many millions at present) while keeping the power dissipated in the chip to an acceptable limit, has made attending to the power dissipated in a logic-gate circuit of paramount importance. Indeed, at the present time, minimizing power dissipation in digital ICs is perhaps the most important design challenge.

In this section, we look at sources of power consumption in digital CMOS circuits and present some metrics that are used in power optimization.

14.6.1 Sources of Power Dissipation

Let us return to the inverter of Fig. 14.17, which dissipates no power when v_I is low and the switch is open. In the other state, however, the power dissipation is approximately V_{DD}^2/R and can be substantial, as we saw in Examples 14.2 and 14.3. This power dissipation occurs even if the inverter is not switching and is thus known as **static power dissipation**.

Another inverter we studied earlier (see Fig. 14.18), which is the basis for the CMOS inverter, exhibits no static power dissipation, a definite advantage. Unfortunately, however, another component of power dissipation arises when a capacitance exists between the output node of the inverter and ground. As we have already seen, this is always the case, for the devices that implement the switches have internal capacitances, the wires that connect the inverter output to other circuits have capacitance, and, of course, there is the input capacitance of whatever circuit the inverter is driving. Now, as the inverter is switched from one state to another, current must flow through the switch(es) to charge (and discharge) the load capacitance. These currents give rise to power dissipation in the switches, called **dynamic power dissipation**.

An expression for the dynamic power dissipation of the inverter of Fig. 14.18 can be derived as follows. Consider first the situation when v_I goes low. The pull-down switch P_D turns off and the pull-up switch P_U turns on. In this state, the inverter can be represented by the equivalent circuit shown in Fig. 14.38(a). Capacitor C will charge through the on-resistance of the pull-up switch, and the voltage across C will increase from 0 to V_{DD}. Denoting by $i_D(t)$ the charging current supplied by V_{DD}, we can write for the instantaneous power drawn from

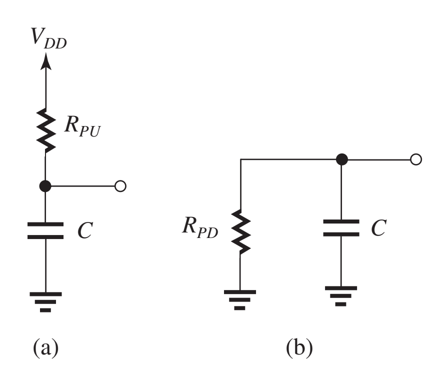


Figure 14.38 Equivalent circuits for calculating the dynamic power dissipation of the inverter in Figure 14.18: **(a)** when v_I is low; **(b)** when v_I is high.

V_{DD} the expression

$$p_{DD}(t) = V_{DD} i_D(t)$$

The energy delivered by the power supply to charge the capacitor can be determined by integrating $p_{DD}(t)$ over the charging interval T_c,

$$E_{DD} = \int_0^{T_c} V_{DD} i_D(t) dt$$

$$= V_{DD} \int_0^{T_c} i_D(t) dt$$

$$= V_{DD} Q$$

where Q is the charge delivered to the capacitor during the charging interval. Since the initial charge on C was zero,

$$Q = CV_{DD}$$

Thus,

$$E_{DD} = CV_{DD}^2 \tag{14.73}$$

Since at the end of the charging process the energy stored on the capacitor is

$$E_{\text{stored}} = \frac{1}{2} CV_{DD}^2 \tag{14.74}$$

we can find the energy dissipated in the pull-up switch as

$$E_{\text{dissipated}} = E_{DD} - E_{\text{stored}} = \frac{1}{2} CV_{DD}^2 \tag{14.75}$$

This energy is dissipated in the on-resistance of switch P_U and is converted to heat.

Next consider the situation when v_I goes high. The pull-up switch P_U turns off and the pull-down switch P_D turns on. The equivalent circuit in this case is that shown in Fig. 14.38(b). Capacitor C is discharged through the on-resistance of the pull-down switch, and its voltage changes from V_{DD} to 0. At the end of the discharge interval, there will be no energy left on the capacitor. Thus all of the energy initially stored on the capacitor, $\frac{1}{2}CV_{DD}^2$, will be dissipated in the pull-down switch,

$$E_{\text{dissipated}} = \frac{1}{2} CV_{DD}^2 \tag{14.76}$$

This amount of energy is dissipated in the on-resistance of switch P_D and is converted to heat.

Thus in each cycle of inverter switching, an amount of energy of $\frac{1}{2}CV_{DD}^2$ is dissipated in the pull-up switch and $\frac{1}{2}CV_{DD}^2$ is dissipated in the pull-down switch, for a total energy loss per cycle of

$$E_{\text{dissipated}}/\text{cycle} = CV_{DD}^2 \tag{14.77}$$

If the inverter is switched at a frequency of f Hz, the dynamic power dissipation of the inverter will be

$$P_{\text{dyn}} = fCV_{DD}^2 \tag{14.78}$$

This is a general expression that does not depend on the inverter circuit details or the values of the on-resistance of the switches.

The expression in Eq. (14.78) indicates that to minimize the dynamic power dissipation, one must strive to reduce the value of C. However, in many cases C is largely determined by the transistors of the inverter itself and cannot be substantially reduced. Another important factor in determining the dynamic power dissipation is the power-supply voltage V_{DD}. Reducing V_{DD} reduces P_{dyn} significantly. This has been a major motivating factor behind the reduction of V_{DD} with every technology generation (see Appendix K). Thus, while the 0.5-μm CMOS process utilized a 5-V power supply, the power-supply voltage used with the 0.13-μm process is only 1.2 V.

Finally, since P_{dyn} is proportional to the operating frequency f, one may be tempted to reduce P_{dyn} by reducing f. However, this is not a viable proposition in light of the desire to operate digital systems at increasingly higher speeds. These newer chips, however, pack much more circuitry on the chip (as many as 2.75 billion transistors) and operate at higher frequencies (microprocessor clock frequencies above 5 GHz are now available). The dynamic power dissipation of such high-density chips can be over 100 W.

In addition to the dynamic power dissipation that results from the periodic charging and discharging of the inverter load capacitance, there is another component of power dissipation in the CMOS inverter that results from the current that flows through Q_P and Q_N during every switching event. Figure 14.39 shows this inverter current as a function of the input voltage

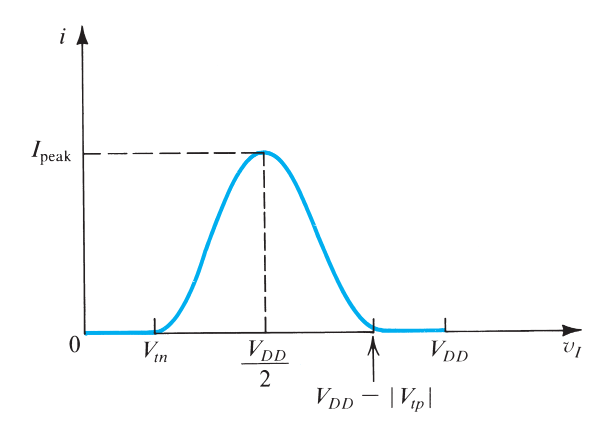


Figure 14.39 The current in the CMOS inverter versus the input voltage.

v_I for a matched inverter. We note that the current peaks at $V_M = V_{DD}/2$. Since at this voltage both Q_N and Q_P operate in saturation, the peak current is given by

$$I_{\text{peak}} = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_n\left(\frac{V_{DD}}{2} - V_{tn}\right)^2 \tag{14.79}$$

The width of the current pulse will depend on the rate of change of v_I with time; the slower the rising edge of the input waveform, the wider the current pulse and the greater the energy drawn from the supply. In general, however, this power component is usually much smaller than P_{dyn}.

EXERCISES

14.19 Find the dynamic power dissipation of the inverter analyzed in Example 14.7 when operated at a 1-GHz frequency. Recall that $C = 6.25$ fF and $V_{DD} = 2.5$ V.
Ans. 39 μW.

14.20 Find the dynamic power dissipation of a CMOS inverter operated from a 1.8-V supply and having a load capacitance of 100 fF. Let the inverter be switched at 100 MHz.
Ans. 32.4 μW

14.21 A particular inverter circuit initially designed in a 0.5-μm process is fabricated in a 0.13-μm process. Assuming that the capacitance C will scale down in proportion to the minimum feature size (more on this in the next chapter) and that the power supply will be reduced from 5 V to 1.2 V, by what factor do you expect the dynamic power dissipation to decrease? Assume that the switching frequency f remains unchanged.
Ans. 66.8

14.6.2 Power–Delay and Energy–Delay Products

One is usually interested in high-speed operation (low t_P) combined with low power dissipation. Unfortunately, these two requirements are often in conflict: Generally, if the designer of an inverter attempts to reduce power dissipation by, say, decreasing the supply voltage V_{DD}, or the supply current, or both, the current-driving capability of the inverter decreases. This in turn results in longer times to charge and discharge the load and parasitic capacitances, and thus the propagation delay increases. It follows that a figure of merit for comparing logic-circuit technologies is the **power–delay product** (PDP) of the basic inverter of the given technology, defined as

$$PDP \equiv P_D t_P \tag{14.80}$$

where P_D is the power dissipation of the inverter. Note that the PDP is an energy quantity and has the units of joules. The lower the PDP, the more effective the inverter and the logic circuits based on the inverter are.

For CMOS logic circuits, the static power dissipation of the inverter is zero,[6] and thus P_D is equal to P_{dyn} and given by Eq. (14.78),

$$P_D = fCV_{DD}^2$$

Thus for the CMOS inverter,

$$PDP = fCV_{DD}^2 t_P \tag{14.81}$$

If the inverter is operated at its theoretical maximum switching speed given by Eq. (14.43), then

$$PDP = \frac{1}{2}CV_{DD}^2 \tag{14.82}$$

From our earlier discussion of dynamic power dissipation we know that $\frac{1}{2}CV_{DD}^2$ is the amount of energy dissipated during each charging or discharging event of the capacitor, that is, for each output transition of the inverter. Thus, the *PDP* has an interesting physical interpretation: *It is the energy consumed by the inverter for each output transition.*

Although the *PDP* is a valuable metric for comparing different technologies for implementing inverters, it is *not* useful as a design parameter for optimizing a given inverter circuit. To appreciate this point, observe that the expression in Eq. (14.82) indicates that the *PDP* can be minimized by reducing V_{DD} as much as possible while, of course, maintaining proper circuit operation. This, however, would not necessarily result in optimal performance, for t_P will increase as V_{DD} is reduced. The problem is that the *PDP* expression in Eq. (14.82) does not in fact have information about t_P. It follows that a better metric can be obtained by multiplying the energy per transition by the propagation delay. We can thus define the **energy–delay product** *EDP* as

$$EDP \equiv \text{Energy per transition} \times t_P$$

$$= \frac{1}{2}CV_{DD}^2 t_P \tag{14.83}$$

We will utilize the *EDP* in later sections.

EXERCISE

14.22 For the CMOS inverter analyzed in Example 14.7, it was found that $C = 6.25$ fF, $V_{DD} = 2.5$ V, and $t_P = 28.7$ ps. Find the power–delay product when the inverter is operated at its theoretical maximum possible operating frequency. Also find *EDP*.
Ans. 19.5 fJ; 5.6×10^{-25} J · s.

[6]The exception to this statement is the power dissipation due to leakage currents and subthreshold conduction in the MOSFETs, discussed in Section 15.1.4.

Summary

- A CMOS logic gate consists of an NMOS pull-down network (PDN) and a PMOS pull-up network (PUN). The PDN conducts for every input combination that requires a low output. Since an NMOS transistor conducts when its input is high, the PDN is most directly synthesized from the expression for the low output $(\overline{Y})$ as a function of the uncomplemented inputs. In a complementary fashion, the PUN conducts for every input combination that corresponds to a high output. Since a PMOS conducts when its input is low, the PUN is most directly synthesized from the expression for a high output (Y) as a function of the complemented inputs.

- The digital logic inverter is the basic building block of digital circuits, just as the amplifier is the basic building block of analog circuits.

- The static operation of a logic inverter is described by its voltage-transfer characteristic (VTC). The VTC determines the inverter noise margins; refer to Fig. 14.13, Fig. 14.15, and to Table 14.1 for the definitions of important VTC points and the noise margins. In particular, note that $NM_H = V_{OH} - V_{IH}$ and $NM_L = V_{IL} - V_{OL}$, and refer to the ideal VTC in Fig. 14.16.

- The inverter is implemented using transistors operating as voltage-controlled switches. There are three possible arrangements, shown in Figs. 14.17, 14.18, and 14.19. The arrangement in Fig. 14.18 results in a high-performance inverter and is the basis for the CMOS inverter studied in Section 14.3.

- The speed of operation of the inverter is characterized by its propagation delay, t_P. Refer to Fig. 14.29 for the definitions of t_{PLH} and t_{PHL}, and note that $t_P = \frac{1}{2}(t_{PLH} + t_{PHL})$.

- Digital ICs usually utilize the minimum channel length of the technology available. Thus for the CMOS inverter, Q_N and Q_P have $L = L_{min}$. If matching is desired, W_p/W_n is selected equal to μ_n/μ_p at the expense of increased area and capacitance. For minimum area, $W_p = W_n$. Also, a frequently used compromise is $W_p = 2W_n$.

- For minimum area, $(W/L)_n$ is selected equal to 1. However, to reduce t_P especially when a major part of C is extrinsic to the inverter, $(W/L)_n$ and correspondingly $(W/L)_p$ can be increased.

- CMOS logic circuits are usually designed to provide equal current-driving capability in both directions. Furthermore, the worst-case values of the pull-up and pull-down currents are made equal to those of the basic inverter. Transistor sizing is based on this principle and makes use of the equivalent W/L ratios of series and parallel devices (Eqs. 14.66 and 14.67).

- An important performance parameter of the inverter is the amount of power it dissipates. There are two components of power dissipation: static and dynamic. The first is the result of current flow in either the 0 or 1 state or both. The second occurs when the inverter is switched and has a capacitor load C. Dynamic power dissipation $P_{\text{dyn}} = fCV_{DD}^2$.

- A metric that combines speed of operation and power dissipation is the power–delay product, $PDP = P_D t_P$. The lower the PDP, the more effective the logic-circuit family is. If dynamic power is dominant, such as in CMOS, the delay-power product for an inverter operated at its theoretical maximum switching frequency is $PDP = \frac{1}{2}CV_{DD}^2$, which is the energy drawn from the supply for a 0-to-1 and a 1-to-0 transition.

- Besides speed of operation and power dissipation, the silicon area required for an inverter is the third significant metric in digital IC design.

- Predominantly because of its low power dissipation and because of its scalability, CMOS is by far the most dominant technology for digital IC design. This situation is expected to continue for many years to come.

- Table 14.2 provides a summary of the important characteristics of the CMOS inverter.

Table 14.2 Summary of Important Characteristics of the CMOS Logic Inverter

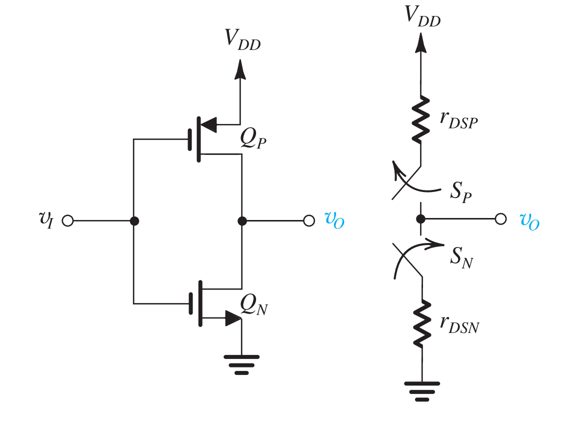
Inverter Output Resistance

- When v_O is low (current sinking):

$$r_{DSN} = 1 \Big/ \left[k_n' \left(\frac{W}{L}\right)_n (V_{DD} - V_{tn}) \right]$$

- When v_O is high (current sourcing):

$$r_{DSP} = 1 \Big/ \left[k_p' \left(\frac{W}{L}\right)_p (V_{DD} - |V_{tp}|) \right]$$



Inverter VTC and Noise Margins

$$V_M = \frac{r(V_{DD} - |V_{tp}|) + V_{tn}}{1 + r} \quad \text{where} \quad r = \sqrt{\frac{k_p'(W/L)_p}{k_n'(W/L)_n}}$$

For matched devices, that is, $\mu_n \left(\frac{W}{L}\right)_n = \mu_p \left(\frac{W}{L}\right)_p$, and $V_{tn} = -V_{tp} = V_t$

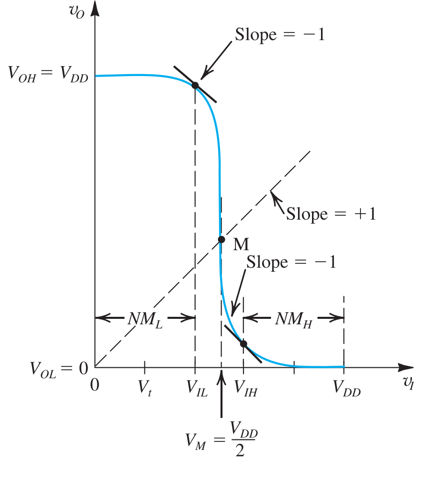
$r = 1$

$V_M = V_{DD}/2$

$V_{IL} = \frac{1}{8}(3V_{DD} + 2V_t)$

$V_{IH} = \frac{1}{8}(5V_{DD} - 2V_t)$

$NM_H = NM_L = \frac{1}{8}(3V_{DD} + 2V_t)$



Propagation Delay (Fig. 14.30)

Using average currents:

$$t_{PHL} \simeq \frac{\alpha_n C}{k_n'(W/L)_n V_{DD}} \quad \text{where} \quad \alpha_n = 2 \Big/ \left[\frac{7}{4} - \frac{3V_{tn}}{V_{DD}} + \left(\frac{V_{tn}}{V_{DD}}\right)^2 \right]$$

$$t_{PLH} \simeq \frac{\alpha_p C}{k_p'(W/L)_p V_{DD}} \quad \text{where} \quad \alpha_p = 2 \Big/ \left[\frac{7}{4} - \frac{3|V_{tp}|}{V_{DD}} + \left(\frac{|V_{tp}|}{V_{DD}}\right)^2 \right]$$

Using equivalent resistances (Fig. 14.31):

$$t_{PHL} = 0.69 R_N C \quad \text{where} \quad R_N = \frac{12.5}{(W/L)_n} \text{ k}\Omega$$

$$t_{PLH} = 0.69 R_P C \quad \text{where} \quad R_P = \frac{30}{(W/L)_p} \text{ k}\Omega$$

For a ramp-input signal, $t_{PHL} \simeq R_N C$ and $t_{PLH} \simeq R_P C$.

Computer Simulation Problems

SIM Problems identified by the Multisim/PSpice icon are intended to demonstrate the value of using SPICE simulation to verify hand analysis and design, and to investigate important issues such as gate noise margins and propagation delays. Instructions to assist in setting up PSpice and Multisim simulations for all the indicated problems can be found in the corresponding files on the website. Note that if a particular parameter value is not specified in the problem statement, you are to make a reasonable assumption.

Section 14.1: CMOS Logic-Gate Circuits

D 14.1 Consider MOS transistors fabricated in a 65-nm process for which $\mu_n C_{ox} = 470\ \mu\text{A/V}^2$, $\mu_p C_{ox} = 190\ \mu\text{A/V}^2$, $V_{tn} = -V_{tp} = 0.35$ V, and $V_{DD} = 1$ V.

(a) Find R_{on} of an NMOS transistor with $W/L = 1.5$.
(b) Find R_{on} of a PMOS transistor with $W/L = 1.5$.
(c) If R_{on} of the PMOS device is to be equal to that of the NMOS device in (a), what must $(W/L)_p$ be?

D 14.2 The CMOS inverter of Fig. 14.2(b) is implemented in a 0.13-μm process for which $\mu_n C_{ox} = 500\ \mu\text{A/V}^2$, $\mu_p C_{ox} = 125\ \mu\text{A/V}^2$, $V_{tn} = -V_{tp} = 0.4$ V, and $V_{DD} = 1.2$ V. The NMOS transistor has $(W/L)_n = 1.5$.

(a) What must $(W/L)_p$ be if Q_N and Q_P are to have equal R_{on} resistances?
(b) Find the value of R_{on}.

D 14.3 Give the CMOS circuit that realizes a three-input NOR gate.

D 14.4 Give the CMOS circuit for a three-input NAND gate.

D 14.5 Find the PUN that corresponds to the PDN shown in Fig. P14.5, and hence the complete CMOS logic circuit. What is the Boolean function realized?

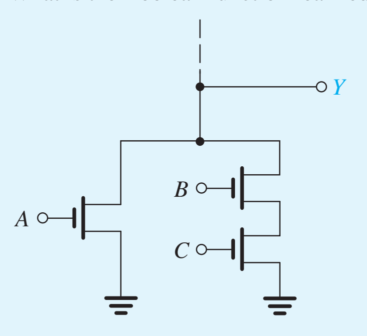


Figure P14.5

D 14.6 Find the PUN that corresponds to the PDN shown in Fig. P14.6, and hence the complete CMOS logic circuit. What is the Boolean function realized?

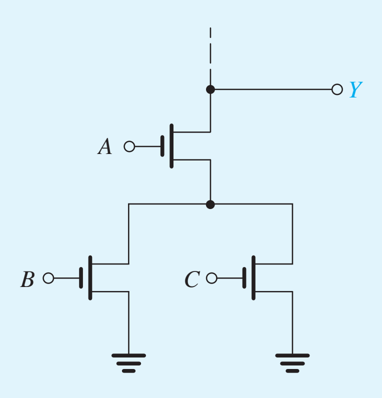


Figure P14.6

D 14.7 Find the PDN that corresponds to the PUN shown in Fig. P14.7, and hence the complete CMOS logic circuit. What is the Boolean function realized?

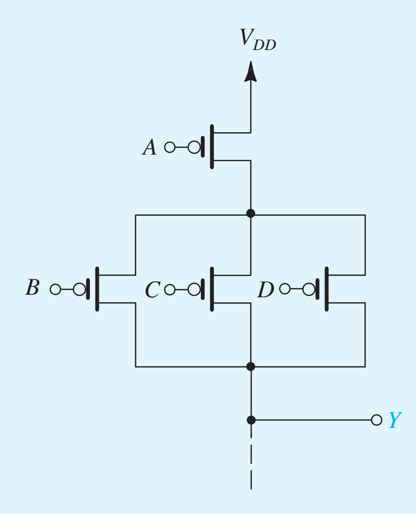


Figure P14.7

D 14.8 Give the CMOS realization for the Boolean function

$$Y = \overline{(A + B)(C + D)}$$

D 14.9 Find the PDN that is the dual of the PUN in Fig. 14.10(a) and hence give a CMOS realization of the exclusive-OR (XOR) function.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

D 14.10 Provide a CMOS logic gate circuit that realizes the function

$$Y = \overline{A}BC + A\overline{B}C + AB\overline{C}$$

How many transistors are required? Explore the possibility of reducing the number of the transistors required.

D 14.11 Sketch a CMOS logic circuit that realizes the function $Y = AB + \overline{A}\,\overline{B}$. This is called the **equivalence** or **coincidence function**.

D 14.12 Sketch a CMOS logic circuit that realizes the function $Y = ABC + \overline{A}\,\overline{B}\,\overline{C}$.

D 14.13 It is required to design a CMOS logic circuit that realizes a three-input, even-parity checker. Specifically, the output Y is to be low when an even number (0 or 2) of the inputs A, B, and C are high.

(a) Give the Boolean function $\overline{Y}$.
(b) Sketch a PDN directly from the expression for $\overline{Y}$. Note that it requires 12 transistors in addition to those in the inverters.
(c) From inspection of the PDN circuit, reduce the number of transistors to 10 (not counting those in the inverters).
(d) Find the PUN as a dual of the PDN in (c), and hence the complete realization.

D 14.14 Give a CMOS logic circuit that realizes the function of a three-input, odd-parity checker. Specifically, the output is to be high when an odd number (1 or 3) of the inputs are high. Attempt a design with 10 transistors (not counting those in the inverters) in each of the PUN and the PDN.

D 14.15 Design a CMOS full-adder circuit with inputs A, B, and C, and two outputs S and C_0 such that S is 1 if one or three inputs are 1, and C_0 is 1 if two or more inputs are 1.

Section 14.2: Digital Logic Inverters

14.16 A particular logic inverter is specified to have $V_{IL} = 0.9$ V, $V_{IH} = 1.2$ V, $V_{OL} = 0.2$ V, and $V_{OH} = 1.8$ V. Find the high and low noise margins, NM_H and NM_L.

14.17 The voltage-transfer characteristic of a particular logic inverter is modeled by three straight-line segments in the manner shown in Fig. 14.13. If $V_{IL} = 1.2$ V, $V_{IH} = 1.3$ V, $V_{OL} = 0.4$ V, and $V_{OH} = 1.8$ V, find:

(a) the noise margins
(b) the value of V_M
(c) the voltage gain in the transition region

14.18 For a particular inverter design using a power supply V_{DD}, $V_{OL} = 0.1V_{DD}$, $V_{OH} = 0.8V_{DD}$, $V_{IL} = 0.4V_{DD}$, and $V_{IH} = 0.6\,V_{DD}$. What are the noise margins? What is the width of the transition region? For a minimum noise margin of 0.4 V, what value of V_{DD} is required?

14.19 A logic-circuit family that used to be very popular is transistor–transistor logic (TTL). The TTL logic gates and other building blocks are available commercially in small-scale-integrated (SSI) and medium-scale-integrated (MSI) packages. Such packages can be assembled on printed-circuit boards to implement a digital system. The device data sheets provide the following specifications of the basic TTL inverter (of the SN7400 type):

Logic-1 input level required to ensure a logic-0 level at the output: MIN (minimum) 2 V
Logic-0 input level required to ensure a logic-1 level at the output: MAX (maximum) 0.8 V
Logic-1 output voltage: MIN 2.4 V, TYP (typical) 3.3 V
Logic-0 output voltage: TYP 0.22 V, MAX 0.4 V
Logic-0-level supply current: TYP 3 mA, MAX 5 mA
Logic-1-level supply current: TYP 1 mA, MAX 2 mA

(a) Find the worst-case values of the noise margins.
(b) Assuming that the inverter is in the logic-1 state 50% of the time and in the logic-0 state 50% of the time, find the average power dissipation in a typical circuit. The power supply is 5 V.

14.20 Consider an inverter implemented as in Fig. 14.17(a). Let $V_{DD} = 2.5$ V, $R = 2\,\text{k}\Omega$, $R_{\text{on}} = 100\,\Omega$, $V_{IL} = 0.8$ V, and $V_{IH} = 1$ V.

(a) Find V_{OL}, V_{OH}, NM_H, and NM_L.
(b) The inverter is driving N identical inverters. Each of these load inverters, or **fan-out** inverters as they are usually called, is specified to require an input current of 0.2 mA when the input voltage (of the fan-out inverter) is high and zero current when the input voltage is low. Noting that the input currents of the fan-out inverters will have to be supplied through R of the driving inverter, find the resulting value of V_{OH} and of NM_H as a function of the number of fan-out inverters N. Hence find the maximum value N can have while the inverter is still providing an NM_H value approximately equal to its NM_L.

(c) Find the power dissipation in the inverter in the two cases: (i) the output is low, and (ii) the output is high and driving the maximum fan-out found in (b).

14.21 For an inverter employing a 2-V supply, suggest an ideal set of values for V_M, V_{IL}, V_{IH}, V_{OL}, V_{OH}, NM_L, NM_H. Also, sketch the VTC. What value of voltage gain in the transition region does your ideal specification imply?

14.22 For a particular inverter, the basic technology used provides an inherent limit to the small-signal, low-frequency voltage gain of 50 V/V. If, with a 2-V supply, the values of V_{OL} and V_{OH} are ideal, but $V_M = 0.4V_{DD}$, what are the best possible values of V_{IL} and V_{IH} that can be expected? What are the best possible noise margins you could expect? Find the large-signal voltage gain, where the gain is defined by $(V_{OH} - V_{OL})/(V_{IL} - V_{IH})$. (*Hint:* Use straight-line approximations for the VTC.)

***14.23** A logic-circuit type intended for use in a digital-signal-processing application in a newly developed hearing aid can operate down to single-cell supply voltages of 1.2 V. If for its inverter, the output signals swing between 0 and V_{DD}, the "gain-of-one" points are separated by less than $\frac{1}{3}$ V_{DD}, and the noise margins are within 30% of one another, what ranges of values of V_{IL}, V_{IH}, V_{OL}, V_{OH}, NM_L, and NM_H can you expect for the lowest possible battery supply?

D 14.24 Design the inverter circuit in Fig. 14.12(a) to provide $V_{OH} = 1.2$ V, $V_{OL} = 50$ mV, and so that the current drawn from the supply in the low-output state is 30 μA. The transistor has $V_t = 0.4$ V, $\mu_n C_{ox} = 500$ μA/V^2, and $\lambda = 0$. Specify the required values of V_{DD}, R_D, and W/L. How much power is drawn from the supply when the output is high? When the output is low?

14.25 For the current-steering circuit in Fig. 14.19, $V_{CC} = 2$ V, $I_{EE} = 0.5$ mA, find the values of R_{C1} and R_{C2} to obtain a voltage swing of 0.5 V at each output. What are the values realized for V_{OH} and V_{OL}?

D 14.26 Refer to the analysis of the resistive-load MOS inverter in Example 14.2 and utilize the expressions derived there for the various inverter parameters. Design the circuit to satisfy the following requirements: $V_{OH} = 1.2$ V, $V_{OL} = 50$ mV, and the power dissipation in the low-output state $= 60$ μW. The transistor available has $V_t = 0.4$ V, $\mu_n C_{ox} = 500$ μA/V^2, and $\lambda = 0$. Specify the required values of V_{DD}, R_D, and W/L. What are the values obtained for V_{IL}, V_M, V_{IH}, NM_L, and NM_H?

D 14.27 Refer to the analysis of the resistive-load MOS inverter in Example 14.2 and utilize the expressions derived there for the various inverter parameters. For a technology for which $V_t = 0.3V_{DD}$, it is required to design the inverter to obtain $V_M = V_{DD}/2$. In terms of V_{DD}, what is the required value of the design parameter V_x? What values are obtained for V_{OH}, V_{OL}, V_{IL}, V_{IH}, NM_H, and NM_L, in terms of V_{DD}? Give numerical values for the case $V_{DD} = 1.2$ V. Now, express the power dissipated in the inverter in its low-output state in terms of the transistor's W/L ratio. Let $k_n' = 500$ μA/V^2. If the power dissipation is to be limited to approximately 100 μW, what W/L ratio is needed and what value of R_D corresponds?

14.28 An earlier form of logic circuits, now obsolete, utilized NMOS transistors only and was appropriately called NMOS logic. The basic inverter, shown in Fig. P14.28, utilizes an NMOS driver transistor Q_1 and another NMOS transistor Q_2, connected as a diode, forms the load of the inverter. Observe that Q_2 operates in saturation at all times. Assume $V_{t1} = V_{t2} = V_t$, $\lambda_1 = \lambda_2 = 0$, and denote $\sqrt{k_{n1}/k_{n2}}$ by k_r. Also neglect the body effect in Q_2 (note that the body of Q_2, not shown, is connected to ground).

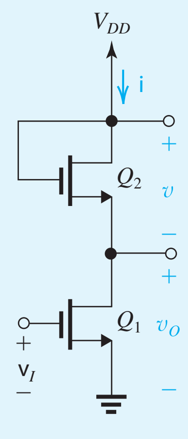


Figure P14.28

(a) Sketch i–v for Q_2 and hence show that for v_I low (i.e., $v_I < V_{tn}$), the output voltage will be $V_{OH} = V_{DD} - V_t$. (*Hint*: Although Q_2 will be conducting zero current, it will have a voltage drop of V_t.)
(b) Taking V_{IL} as the value of v_I at which Q_1 begins to conduct and v_O begins to fall, find V_{IL}.
(c) Find the relationship between v_O and v_I in the transition region. This is the region for which $v_I > V_t$ and both Q_1 and Q_2 are operating in saturation. Show that the relationship is linear and find its slope.
(d) If $V_{OL} \simeq 0$ V, find the current I_{DD} drawn from V_{DD} and hence the average power dissipation in the inverter, assuming that it spends half the time in each of its two states.
(e) Find numerical values for all the parameters asked for above for the case $V_{DD} = 1.8$ V, $V_t = 0.5$ V, $(W/L)_1 = 5$, $(W/L)_2 = \frac{1}{5}$, and $\mu_n C_{ox} = 300\ \mu\text{A/V}^2$.

14.29 For the pseudo-NMOS inverter analyzed in Example 14.3 and in Exercise 14.5, what is the value of r that results in $V_M = V_{DD}/2 = 0.9$ V?

14.30 Repeat Example 14.3 for a pseudo-NMOS inverter fabricated in a 0.13-μm CMOS technology for which $V_{DD} = 1.2$ V, $|V_t| = 0.4$ V, $k_n/k_p = 5$, and $k_n = 500\ \mu\text{A/V}^2$. Find V_{OH}, V_{OL}, I_{DD}, and the average power dissipation P_{av}. Also, use the expression given in Exercise 14.5 to evaluate V_M.

Section 14.3: The CMOS Inverter

14.31 Consider a CMOS inverter fabricated in a 65-nm CMOS process for which $V_{DD} = 1$ V, $V_{tn} = -V_{tp} = 0.35$ V, and $\mu_n C_{ox} = 2.5\mu_p C_{ox} = 470\ \mu\text{A/V}^2$. In addition, Q_N and Q_P have $L = 65$ nm and $(W/L)_n = 1.5$.

(a) Find W_p that results in $V_M = V_{DD}/2$. What is the silicon area utilized by the inverter in this case?
(b) For the matched case in (a), find the values of V_{OH}, V_{OL}, V_{IH}, V_{IL}, NM_L, and NM_H.
(c) For the matched case in (a), find the output resistance of the inverter in each of its two states.

SIM **14.32** Consider a CMOS inverter fabricated in a 0.25-μm CMOS process for which $V_{DD} = 2.5$ V, $V_{tn} = -V_{tp} = 0.5$ V, and $\mu_n C_{ox} = 3.5\ \mu_p C_{ox} = 115\ \mu\text{A/V}^2$. In addition, Q_N and Q_P have $L = 0.25\ \mu$m and $(W/L)_n = 1.5$. Investigate the variation of V_M with the ratio W_p/W_n. Specifically, calculate V_M for (a) $W_p = 3.5W_n$ (the matched case), (b) $W_p = W_n$ (the minimum-size case); and (c) $W_p = 2W_n$ (a compromise case). For cases (b) and (c), estimate the approximate reduction in NM_L and silicon area relative to the matched case (a).

14.33 For a technology in which $V_{tn} = 0.3V_{DD}$, show that the maximum current that the inverter can sink while its low-output level does not exceed 0.1 V_{DD} is $0.065\ k_n'(W/L)_n V_{DD}^2$. For $V_{DD} = 1.3$ V, $k_n' = 500\ \mu\text{A/V}^2$, find $(W/L)_n$ that permits this maximum current to be 0.1 mA.

14.34 A CMOS inverter for which $k_n = 5k_p = 200\ \mu\text{A/V}^2$ and $V_t = 0.5$ V is connected as shown in Fig. P14.34 to a sinusoidal signal source having a Thévenin equivalent voltage of 0.1-V peak amplitude and resistance of 100 kΩ. What signal voltage appears at node A with $v_I = +1.5$ V? With $v_I = -1.5$ V?

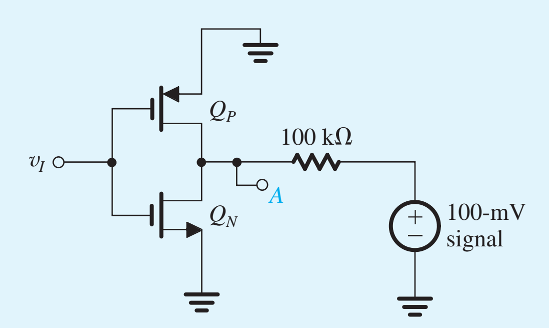


Figure P14.34

D **14.35** There are situations in which Q_N and Q_P of the CMOS inverter are deliberately mismatched to realize a certain desired value for V_M. Show that the value required of the parameter r of Eq. (14.40) is given by

$$r = \frac{V_M - V_{tn}}{V_{DD} - |V_{tp}| - V_M}$$

For a 0.13-μm process characterized by $V_{tn} = -V_{tp} = 0.4$ V, $V_{DD} = 1.3$ V, and $\mu_n = 4\mu_p$, find the ratio W_p/W_n required to obtain $V_M = 0.6V_{DD}$.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

14.36 Consider the CMOS inverter of Fig. 14.22 with Q_N and Q_P matched and with the input v_I rising slowly from 0 to V_{DD}. At what value of v_I does the current flowing through Q_N and Q_P reach its peak? Give an expression for the peak current, neglecting λ_n and λ_p. For $k_n' = 500\ \mu A/V^2$, $(W/L)_n = 1.5$, $V_{DD} = 1.3$ V, and $V_{tn} = 0.4$ V, find the value of the peak current.

14.37 Repeat Example 14.4 for a CMOS inverter fabricated in a 0.13-μm process for which $V_{DD} = 1.3$ V, $V_{tn} = |V_{tp}| = 0.4$ V, $\mu_n = 4\mu_p$, and $\mu_n C_{ox} = 500\ \mu A/V^2$. In addition, Q_N and Q_P have $L = 0.13$ μm and $(W/L)_n = 1.5$. For part (a) use $V_M = V_{DD}/2 = 0.65$ V.

Section 14.4: Dynamic Operation of the CMOS Inverter

14.38 For the circuit shown in Fig. P14.38, let switch S open at $t = 0$.

(a) Give the expression for $v_O(t)$.
(b) For $I = 1$ mA and $C = 10$ pF, find the time at which v_O reaches 1 V.

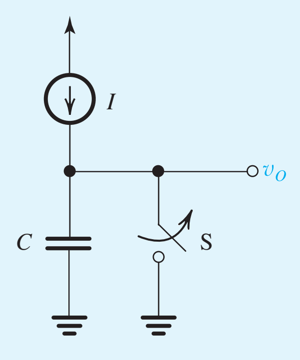


Figure P14.38

14.39 For the circuit in Fig. P14.39, let C be charged to 10 V and switch S closes at $t = 0$.

(a) Give the expression for $v_O(t)$.
(b) For $C = 100$ pF and $R = 1$ kΩ, find t_{PHL} and t_f.

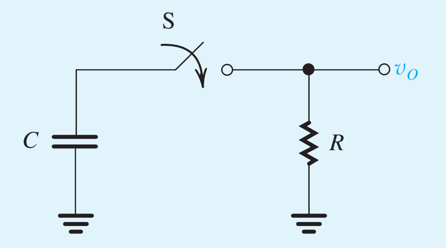


Figure P14.39

14.40 For the inverter circuit in Fig. P14.40, let v_I go from V_{DD} to 0 V at $t = 0$. At $t = 0+$, $v_O = V_{OL}$. Find expressions for V_{OH}, $v_O(t)$, and t_{PLH}. If $R = 10$ kΩ, what is the largest value of C that ensures that t_{PLH} is at most 100 ps?

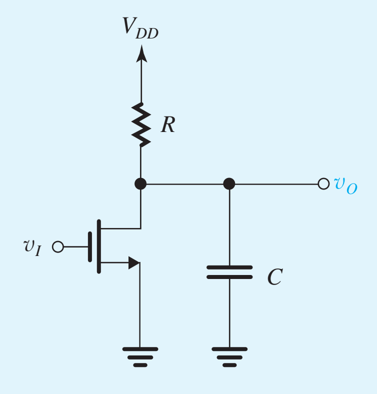


Figure P14.40

14.41 For the inverter of Fig. 14.18(a) with a capacitance C connected between the output and ground, let the on-resistance of P_U be 2 kΩ and that of P_D be 1 kΩ. If the capacitance $C = 50$ fF, find t_{PLH}, t_{PHL}, and t_P.

14.42 A logic inverter is implemented using the arrangement of Fig. 14.18 with switches having $R_{on} = 2$ kΩ, $V_{DD} = 1.8$ V, and $V_{IL} = V_{IH} = V_{DD}/2$.

(a) Find V_{OL}, V_{OH}, NM_L, and NM_H.
(b) If v_I rises instantaneously from 0 V to +1.8 V and assuming the switches operate instantaneously—that is, at $t = 0$, PU opens and PD closes—find an expression for $v_O(t)$, assuming that a capacitance C is connected between the output node and ground. Hence find the high-to-low

propagation delay (t_{PHL}) for $C = 0.1$ pF. Also find t_{THL} (see Fig. 14.29).

(c) Repeat (b) for v_I falling instantaneously from +1.8 V to 0 V. Again assume that PD opens and PU closes instantaneously. Find an expression for $v_O(t)$, and hence find t_{PLH} and t_{TLH}.

14.43 In a particular logic family, the standard inverter, when loaded by a similar circuit, has a propagation delay specified to be 0.9 ns:

(a) If the current available to charge a load capacitance is half as large as that available to discharge the capacitance, what do you expect t_{PLH} and t_{PHL} to be?

(b) If when an external capacitive load of 0.5 pF is added at the inverter output, its propagation delays increase by 50%, what do you estimate the normal combined capacitance of inverter output and input to be?

(c) If without the additional 0.5-pF load connected, the load inverter is removed and the propagation delays were observed to decrease by 40%, estimate the two components of the capacitance found in (b): that is, the component due to the inverter output and other associated parasitics, and the component due to the input of the load inverter.

***14.44** Consider an inverter for which t_{PLH}, t_{PHL}, t_{TLH}, and t_{THL} are 20 ns, 10 ns, 30 ns, and 15 ns, respectively. The rising and falling edges of the inverter output can be approximated by linear ramps. Also, for simplicity, we define t_{TLH} to be 0% to 100% (rather than 10% to 90%) rise time, and similarly for t_{THL}. Two such inverters are connected in tandem and driven by an ideal input having zero rise and fall times. Calculate the time taken for the output voltage to complete its excursion for (a) a rising input and (b) a falling input. What is the propagation delay for the inverter?

SIM **14.45** For a CMOS inverter fabricated in a 0.13-μm process with $V_{DD} = 1.2$ V, $V_{tn} = -V_{tp} = 0.4$ V, $k_n' = 4k_p' = 430\ \mu\text{A/V}^2$, and having $(W/L)_n = 1.5$ and $(W/L)_p = 3$, find t_{PHL}, t_{PLH}, and t_P when the equivalent load capacitance $C = 10$ fF. Use the method of average currents.

D **14.46** Consider a matched CMOS inverter fabricated in the 0.13-μm process specified in Problem 14.45. If $C = 30$ fF, use the method of average currents to determine the required (W/L) ratios so that $t_P \leq 80$ ps.

14.47 For the CMOS inverter in Exercise 14.11 use the method of equivalent resistance to determine t_{PHL}, t_{PLH}, and t_P.

14.48 Use the method of equivalent resistance to determine the propagation delay of a minimum-size inverter, that is, one for which $(W/L)_n = (W/L)_p = 1$, designed in a 0.13-μm technology. The equivalent load capacitance $C = 20$ fF.

D **14.49** Use the method of equivalent resistance to design an inverter to be fabricated in a 0.13-μm technology. It is required that for $C = 10$ fF, $t_{PLH} = t_{PHL}$, and $t_P \leq 50$ ps.

14.50 The method of average currents yields smaller values for t_{PHL} and t_{PLH} than those obtained by the method of equivalent resistances. Most of this discrepancy is due to the fact that the formula we derived for I_{av} does not take into account velocity saturation. As will be seen in Section 15.1.2, velocity saturation reduces the current significantly. Using the results in Example 14.6, by what factor do you estimate the current reduction to be in the NMOS transistor? Since t_{PLH} does not change, what do you conclude about the effect of velocity saturation on the PMOS transistor in this technology?

14.51 Use the method of average currents to estimate t_{PHL}, t_{PLH}, and t_P of a CMOS inverter fabricated in a 65-nm process for which $V_{tn} = |V_{tp}| = 0.35$ V, $V_{DD} = 1$ V, $\mu_n C_{ox} = 470\ \mu\text{A/V}^2$, and $\mu_p C_{ox} = 190\ \mu\text{A/V}^2$. The inverter has $(W/L)_n = 1.5$ and $(W/L)_p = 3$, and the total capacitance at the inverter output node is 10 fF. Also, find the theoretical maximum frequency at which this inverter can be operated.

14.52 Find the propagation delay for a minimum-size inverter for which $k_n' = 4k_p' = 380\ \mu\text{A/V}^2$ and $(W/L)_n = (W/L)_p = 0.27\ \mu\text{m}/0.18\ \mu\text{m}$, $V_{DD} = 1.8$ V, $V_{tn} = -V_{tp} = 0.5$ V, and the capacitance is roughly 4 fF/μm of device width plus 2 f F/device. There is an additional load capacitance of 5 fF. What does t_P become if the design is changed to a matched one? Use the method of average current.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

14.53 A matched CMOS inverter fabricated in a process for which $C_{ox} = 3.7$ fF/μm^2, $\mu_n C_{ox} = 180$ μA/V^2, $\mu_p C_{ox} = 45$ μA/V^2, $V_{tn} = -V_{tp} = 0.7$ V, and $V_{DD} = 3.3$ V, uses $W_n = 0.75$ μm and $L_n = L_p = 0.5$ μm. The overlap capacitance and the effective drain–body capacitance per micrometer of gate width are 0.4 fF and 1.0 fF, respectively. The wiring capacitance is $C_w = 2$ fF. If the inverter is driving another identical inverter, find t_{PLH}, t_{PHL}, and t_P. For how much additional capacitance load does the propagation delay increase by 50%?

Section 14.5: Transistor Sizing

14.54 An inverter whose equivalent load capacitance C is composed of 15 fF contributed by the inverter transistors, and 45 fF contributed by the wiring and other external circuitry, has been found to have a propagation delay of 80 ps. By what factor must $(W/L)_n$ and $(W/L)_p$ be increased so as to reduce t_P to 40 ps? By what factor is the inverter area increased?

D *14.55 In this problem we investigate the effect of the selection of the ratio W_p/W_n on the propagation delay of an inverter driving an identical inverter, as in Fig. 14.32. Assume all transistors have the same L.

(a) Noting that except for C_w each of the capacitances in Eqs. (14.58) and (14.59) is proportional to the width of the relevant transistor, show that C can be expressed as

$$C = C_n\left(1 + \frac{W_p}{W_n}\right) + C_w$$

where C_n is determined by the NMOS transistors.

(b) Using the equivalent resistances R_N and R_P, show that for $(W/L)_n = 1$,

$$t_{PHL} = 8.625 \times 10^3 C$$

$$t_{PLH} = \frac{20.7 \times 10^3}{W_p/W_n} C$$

(c) Use the results of (a) and (b) to determine t_P in the case $W_p = W_n$, in terms of C_n and C_w.

(d) Use the results of (a) and (b) to determine t_P in the matched case: that is, when W_p/W_n is selected to yield $t_{PHL} = t_{PLH}$.

(e) Compare the t_P values in (c) and (d) for the two extreme cases:

(i) $C_w = 0$

(ii) $C_w \gg C_n$

What do you conclude about the selection of W_p/W_n?

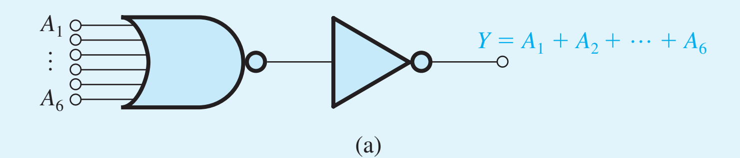
D 14.56 Consider the CMOS gate shown in Fig. 14.9. Specify W/L ratios for all transistors in terms of the ratios n and p of the basic inverter, such that the worst-case t_{PHL} and t_{PLH} of the gate are equal to those of the basic inverter.

D 14.57 Find appropriate sizes for the transistors used in the exclusive-OR circuit of Fig. 14.10(b). Assume that the basic inverter has $(W/L)_n = 0.20$ μm/0.13 μm and $(W/L)_p =$ 0.40 μm/0.13 μm. What is the total area, including that of the required inverters?

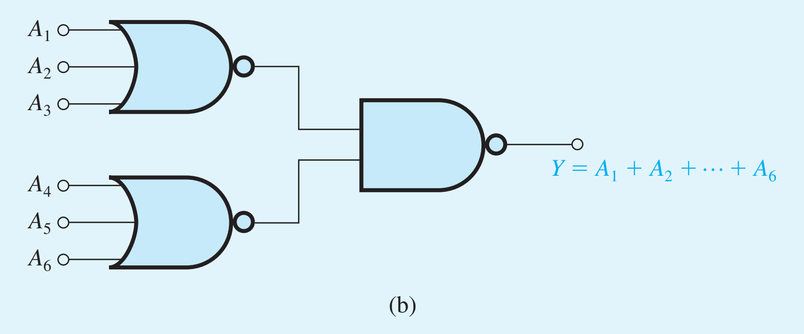
14.58 Consider a four-input CMOS NAND gate for which the transient response is dominated by a fixed-size capacitance between the output node and ground. Compare the values of t_{PLH} and t_{PHL}, obtained when the devices are sized as in Fig. 14.35, to the values obtained when all n-channel devices have $W/L = n$ and all p-channel devices have $W/L = p$.

14.59 Figure P14.59 shows two approaches to realizing the OR function of six input variables. The circuit in Fig. P14.59(b), though it uses additional transistors, has in fact less total area and lower propagation delay because it uses NOR gates with lower fan-in. Assuming that the transistors in both circuits are properly sized to provide each gate with a current-driving capability equal to that of the basic matched inverter, find the number of transistors and the total area of each circuit. Assume the basic inverter to have a $(W/L)_n$ ratio of 0.20 μm/0.13 μm and a $(W/L)_p$ ratio of 0.40 μm/0.13 μm.

***14.60** Consider the two-input CMOS NOR gate of Fig. 14.7 whose transistors are properly sized so that the current-driving



(a)



(b)

Figure P14.59

capability in each direction is equal to that of a matched inverter. For $|V_t| = 1$ V and $V_{DD} = 5$ V, find the gate threshold in the cases for which (a) input terminal A is connected to ground and (b) the two input terminals are tied together. Neglect the body effect in Q_{PB}.

14.61 A chain of four inverters whose sizes are scaled by a factor x is used to drive a load capacitance $C_L = 1200\,C$, where C is the input capacitance of the standard inverter (which is the first in the chain).

(a) Without increasing the number of inverters in the chain, find the optimum value of x that results in minimizing the overall delay t_P and find the resulting value of t_P in terms of the time constant CR, where R is the output resistance of the standard inverter.
(b) If you are allowed to increase the number of inverters in the chain, what is the number of inverters and the value of x that result in minimizing the total path delay t_P? What is the value of t_P achieved?

14.62 The purpose of this problem is to find the values of n and x that result in minimum path delay t_P for the inverter chain in Fig. 14.37(c).

(a) Show that

$$t_P = \tau_{\text{total}} = (n-1)xRC + \frac{1}{x^{n-1}}RC_L$$

(b) Differentiate the expression for t_P in (a) relative to x and set the derivative to zero. Thus show that the first condition for optimality is

$$x^n = \frac{C_L}{C}$$

(c) Differentiate the expression for t_P in (a) relative to n and set the derivative to zero. Thus show that the second condition for optimality is

$$x^n\left(\frac{C}{C_L}\right) = \ln x$$

(d) Combine the expressions in (b) and (c) to show that the value of x for minimum overall delay is

$$x = e$$

Section 14.6: Power Dissipation

14.63 An IC inverter fabricated in a 0.18-μm CMOS process is found to have a load capacitance of 10 fF. If the inverter is operated from a 1.8-V power supply, find the energy needed to charge and discharge the load capacitance. If the IC chip has 2 million of these inverters operating at an average switching frequency of 1 GHz, what is the power dissipated in the chip? What is the average current drawn from the power supply?

14.64 Consider a logic inverter of the type shown in Fig. 14.18. Let $V_{DD} = 1$ V, and let a 5-fF capacitance be connected between the output node and ground. If the inverter is switched at the rate of 2 GHz, determine the dynamic power dissipation. What is the average current drawn from the dc power supply?

14.65 In a particular logic-circuit technology, operating with a 3.3-V supply, the basic inverter draws (from the supply) a current of 60 μA in one state and 0 μA in the other. When the inverter is switched at the rate of 100 MHz, the average supply current becomes 150 μA. Estimate the equivalent capacitance at the output node of the inverter.

14.66 A collection of logic gates for which the static power dissipation is zero, and the dynamic power dissipation is 10 mW is operating at 50 MHz with a 5-V supply. By what fraction could the power dissipation be reduced if operation at 3.3 V were possible? If the frequency of operation is reduced by the same factor as the supply voltage (i.e., 3.3/5), what *additional* power can be saved?

14.67 A particular logic gate has t_{PLH} and t_{PHL} of 30 ns and 50 ns, respectively, and dissipates 1 mW with output low and 0.6 mW with output high. Calculate the corresponding delay–power product (under the assumption of a 50% duty-cycle signal and neglecting dynamic power dissipation).

D *14.68 We wish to investigate the design of the inverter shown in Fig. 14.17(a). In particular, we wish to determine the value for R. Selection of a suitable value for R is determined by two considerations: propagation delay and power dissipation.

(a) Show that if v_I changes instantaneously from high to low and assuming that the switch opens instantaneously, the output voltage obtained across a load capacitance C will be

$$v_O(t) = V_{OH} - (V_{OH} - V_{OL})e^{-t/\tau_1}$$

where $\tau_1 = CR$. Hence show that the time required for $v_O(t)$ to reach the 50% point, $\frac{1}{2}(V_{OH} + V_{OL})$, is

$$t_{PLH} = 0.69CR$$

(b) Following a steady state, if v_I goes high and assuming that the switch closes immediately and has the equivalent circuit in Fig. 14.17(c), show that the output falls exponentially according to

$$v_O(t) = V_{OL} + (V_{OH} - V_{OL})e^{-t/\tau_2}$$

where $\tau_2 = C(R \| R_{on}) \simeq CR_{on}$ for $R_{on} \ll R$. Hence show that the time for $v_O(t)$ to reach the 50% point is

$$t_{PHL} = 0.69CR_{on}$$

(c) Use the results of (a) and (b) to obtain the inverter propagation delay, defined as the average of t_{PLH} and t_{PHL} as

$$t_P \simeq 0.35CR \quad \text{for } R_{on} \ll R$$

(d) Show that for an inverter that spends half the time in the logic-0 state and half the time in the logic-1 state, the average static power dissipation is

$$P = \frac{1}{2}\frac{V_{DD}^2}{R}$$

(e) Now that the trade-offs in selecting R should be clear, show that, for $V_{DD} = 5$ V and C = 10 pF, to obtain a propagation delay no greater than 5 ns and a power dissipation no greater than 15 mW, R should be in a specific range. Find that range and select an appropriate value for R. Then determine the resulting values of t_P and P.

D 14.69 A logic-circuit family with zero static power dissipation normally operates at $V_{DD} = 2.5$ V. To reduce its dynamic power dissipation, operation at 1.8 V is considered. It is found, however, that the currents available to charge and discharge load capacitances also decrease. If current is (a) proportional to V_{DD} or (b) proportional to V_{DD}^2, what reductions in maximum operating frequency do you expect in each case? What fractional change in delay–power product do you expect in each case?

14.70 In this problem we estimate the CMOS inverter power dissipation resulting from the current pulse that flows in Q_N and Q_P when the input pulse has finite rise and fall times. Refer to Fig. 14.39 and let $V_{tn} = -V_{tp} = 0.5$ V, $V_{DD} = 1.8$ V, and $k_n = k_p = 450\ \mu A/V^2$. Let the input rising and falling edges be linear ramps with the 0-to-V_{DD} and V_{DD} -to-0 transitions taking 1 ns each. Find I_{peak}. To determine the energy drawn from the supply per transition, assume that the current pulse can be approximated by a triangle with a base corresponding to the time for the rising or falling edge to go from V_t to $V_{DD} - V_t$, and the height equal to I_{peak}. Also, determine the power dissipation that results when the inverter is switched at 100 MHz.

CHAPTER 15

Advanced Topics in Digital Integrated-Circuit Design

IN THIS CHAPTER YOU WILL LEARN

1. The implications of technology scaling (Moore's law) over more than 40 years and continuing, and some of the current challenges in the design of deep-submicron ($L < 0.25\ \mu m$) circuits.
2. How and why CMOS has become the dominant technology for digital IC design.
3. That by replacing the pull-up network (PUN) of a CMOS logic gate by a single PMOS transistor that is permanently on, considerable savings in transistor count and silicon area can be achieved in gates with high fan-in. The resulting circuits are known as pseudo-NMOS.
4. That a useful and conceptually simple form of MOS logic circuit, known as pass-transistor logic (PTL), utilizes MOS transistors as series switches in the signal path from input to output.
5. That a very effective switch for both analog and digital applications, known as transmission gate, is formed by connecting an NMOS and a PMOS transistor in parallel.
6. That eliminating the pull-up network and placing two complementary switches, operated by a clock signal, in series with the pull-down network of a CMOS gate results in an interesting and useful class of circuits known as dynamic logic.
7. How the BJT differential-pair configuration is used as a current switch to realize the fastest commercially available logic-circuit family: emitter-coupled logic (ECL).
8. How the MOSFET and the BJT are combined in BiCMOS circuits in ways that take advantage of the best attributes of each device.

Introduction

In this chapter we study a number of advanced topics in digital logic-circuit design. We begin by taking a closer look at the implications of Moore's law. Specifically, over the past 45 years or so, the MOSFET dimensions have been reduced by a factor of 2 about every five years. This scaling has been accompanied by reductions in V_{DD} and V_t. The opportunities provided and challenges posed by scaling are studied in Section 15.1. We then survey the field of digital IC technologies in order to place CMOS in proper perspective.

Standard CMOS logic, which we studied in Chapter 14, excels in almost every performance category: It is easy to design, has the maximum possible voltage swing, is

robust from a noise-immunity standpoint, dissipates no static power, and can be designed to provide equal high-to-low and low-to-high propagation delays. Its main disadvantage is the requirement of two transistors for each additional gate input, which for gates with high fan-in can make the chip area large and increase the total capacitance and, correspondingly, the propagation delay and the dynamic power dissipation. For this reason designers of digital integrated circuits have been searching for forms of CMOS logic circuits that can be used to supplement standard CMOS. This chapter presents three such forms that reduce the required number of transistors but incur other costs. These forms are not intended to replace standard CMOS, but are rather to be used in special applications for special purposes.

Pseudo-NMOS logic, studied in Section 15.3, replaces the pull-up network (PUN) in a CMOS logic gate by a single permanently "on" PMOS transistor. The reduction in transistor count and silicon area comes at the expense of static power dissipation. As well, the output low-level V_{OL} becomes dependent on the transistors' W/L ratios.

Pass-transistor logic (PTL), studied in Section 15.4, utilizes MOS transistors as switches in the series path from input to output. Though simple and attractive for special applications, PTL does not restore the signal level and thus requires the occasional use of standard CMOS inverters to avoid signal-level degradation, especially in long chains of switches.

The dynamic logic circuits studied in Section 15.5 dispense with the PUN and place two complementary switches in series with the PDN. The switches are operated by a clock, and the gate output is stored on the load capacitance. Here the reduction in transistor count is achieved at the expense of a more complex design that is less robust than static CMOS.

Although CMOS accounts for the vast majority of digital integrated circuits, there is a bipolar logic-circuit family that is still of some interest. This is emitter-coupled logic (ECL), which we study briefly in Section 15.6.1. Finally, in Section 15.6.2 we show how the MOSFET and the BJT can be combined in ways that take advantage of the best properties of each, resulting in what are known as BiCMOS circuits.

The sections of this chapter are almost independent modules, thus selected ones can be studied as they come up, and others may be deferred to a later time.



15.1 Implications of Technology Scaling: Issues in Deep-Submicron Design

As mentioned in Chapter 5, and in a number of locations throughout the book, the minimum MOSFET channel length has been continually reduced over the past 50 years or so. In fact, a new CMOS fabrication technology has been introduced every 2 or 3 years, with the minimum allowable channel length reduced by about 30%, that is, to 0.7 the value in the preceding generation. Thus, with every new **technology generation**, the device area has been reduced by a factor of $1/(0.7 \times 0.7)$ or approximately 2, allowing the fabrication of twice as many devices on a chip of the same area. This astounding phenomenon, predicted nearly 50 years ago by Gordon Moore, has become known as **Moore's law**. It is this ability to pack an exponentially increasing number of transistors on an IC chip that has resulted in the continuing reduction in the cost per logic function.

Figure 15.1 shows the exponential reduction in MOSFET channel length (by a factor of 2 every 5 years) over a 40-year period, with the dots indicating some of the prominent **technology generations**, or **nodes**. Thus, we see the 10-μm process of the early 1970s, the

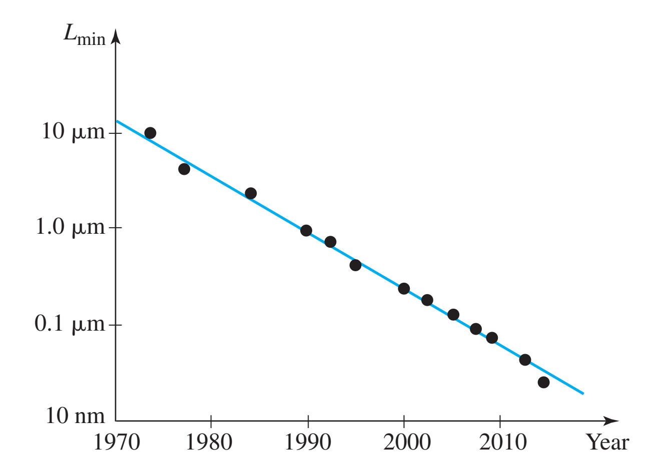


Figure 15.1 The MOSFET channel length has been reduced by a factor of 2 about every 5 years. This phenomenon, known as Moore's law, is continuing.

submicron ($L < 1$ μm) processes of the early 1990s, and the deep-submicron ($L < 0.25$ μm) processes of the last decade, including the current 22-nm process. A microprocessor chip fabricated in a 22-nm CMOS process, clocked at 3.8 GHz and having 4.31 billion transistors, was announced in 2014. Deep-submicron (DSM) processes present the circuit designer with a host of new opportunities and challenges. It is our purpose in this section to briefly consider some of these.

15.1.1 Silicon Area

We begin this section with a brief discussion of silicon area. In addition to minimizing power dissipation and propagation delay, an important objective in the design of digital VLSI circuits is the minimization of silicon area per logic gate. The smaller area requirement enables the fabrication of a larger number of gates per chip, which has economic and space advantages from a system-design standpoint. Area reduction occurs in three different ways: through advances in processing technology that enable the reduction of the minimum device size, through advances in circuit-design techniques, and through careful chip layout. In this book, our interest lies in circuit design, and we shall make frequent comments on the relationship between the design of a circuit and its silicon area. As a general rule, the simpler the circuit, the smaller the area required. As we have seen in Section 14.5, the circuit designer has to decide on device sizes. Choosing smaller devices has the obvious advantage of requiring smaller silicon area and at the same time reducing parasitic capacitances and thus increasing speed. Smaller devices, however, have lower current-driving capability, which tends to increase delay. Thus, as in all engineering design problems, there is a trade-off to be quantified and exercised in a manner that optimizes whatever aspect of the design is thought to be critical for the application at hand.

15.1.2 Scaling Implications

Table 15.1 provides a summary of the implications of scaling the device dimensions by a factor $1/S$, where $S > 1$. As well, we assume that V_{DD} and V_t are scaled by the same factor.

Table 15.1 Implications of Device and Voltage Scaling

	Parameter	Relationship	Scaling Factor
1	W, L, t_{ox}		$1/S$
2	V_{DD}, V_t		$1/S$
3	Area/Device	WL	$1/S^2$
4	C_{ox}	ϵ_{ox}/t_{ox}	S
5	k_n', k_p'	$\mu_n C_{ox}$, $\mu_p C_{ox}$	S
6	C_{gate}	WLC_{ox}	$1/S$
7	t_P (intrinsic)	$\alpha C/k' V_{DD}$	$1/S$
8	Energy/Switching cycle (intrinsic)	CV_{DD}^2	$1/S^3$
9	P_{dyn}	$f_{\max} CV_{DD}^2 = \dfrac{CV_{DD}^2}{2t_P}$	$1/S^2$
10	Power density	P_{dyn}/Device area	1

Although the scaling of V_{DD} has occurred for a number of technology nodes (e.g., from 5 V for the 0.5-μm process down to 1.2 V for the 0.13-μm process and 1 V for the 65-nm process), V_t has been reduced but not by the same factor. Thus the assumption in row 2 of Table 15.1 is not entirely correct. Nevertheless, our interest here is to gain a general appreciation for the effects of scaling.

Table 15.1 provides the relationships for the various transistor and inverter parameters in order to show how the resulting scale factors are obtained. We thus see that the device area scales by $1/S^2$; the oxide capacitance C_{ox}, and the transconductance parameters k_n' and k_p' scale by S; and the MOSFET gate capacitance scales by $1/S$. It is important to note that the component of the inverter propagation delay due to the transistor capacitances (i.e., excluding the wiring capacitance) scales by $1/S$; this very useful result of scaling implies that the circuit can be operated at S times the frequency; that is, the speed of operation increases by a factor S. Equally important, the dynamic power dissipation scales by $1/S^2$. This, of course, is a major motivating factor behind the scaling of V_{DD}. Another motivating factor is the need to keep the electric fields in the MOSFETs within acceptable bounds.

Although the dynamic power dissipation is scaled by $1/S^2$, the power per unit area remains unchanged. Nevertheless, for a number of reasons, as the size and complexity of digital IC chips continue to increase, so does their power dissipation. Indeed power dissipation has now become the number-one issue in IC design. The problem is further exacerbated by the static power dissipation, arising from both subthreshold conduction and diode leakage currents, that plagues deep-submicron CMOS devices. We will discuss this issue shortly.

EXERCISES

15.1 By what factor does the power–delay product *PDP* change if an inverter is fabricated in a 0.13-μm technology rather than a 0.25-μm technology? Assume $S \simeq 2$.
Ans. *PDP* decreases by a factor of 8.

15.2 If V_{DD} and V_t are kept constant, which entries in Table 15.1 change and to what value?
Ans. t_P now scales by $1/S^2$; the energy/switching cycle now scales by $1/S$ only; P_{dyn} now scales by S; and the power density now scales by S^3 (a major problem).

15.1.3 Velocity Saturation

The short channels of MOSFETs fabricated in deep-submicron processes give rise to physical phenomena not present in long-channel devices, and thus to changes in the MOSFET *i–v* characteristics. The most important of these **short-channel** effects is **velocity saturation**. Here we refer to the drift velocity of electrons in the channel of an NMOS transistor (holes in PMOS) under the influence of the longitudinal electric field established by v_{DS}. In our derivation of the MOSFET *i–v* characteristics in Section 5.1, we assumed that the velocity v_n of the electrons in an *n*-channel device is given by

$$v_n = \mu_n E \tag{15.1}$$

where E is the electric field given by

$$E = \frac{v_{DS}}{L} \tag{15.2}$$

The relationship in Eq. (15.1) applies as long as E is below a critical value E_{cr} that falls in the range 1 V/μm to 5 V/μm. For $E > E_{cr}$, the drift velocity saturates at a value v_{sat} of approximately 10^7 cm/s. Figure 15.2 shows a sketch of v_n versus E. Although the change from

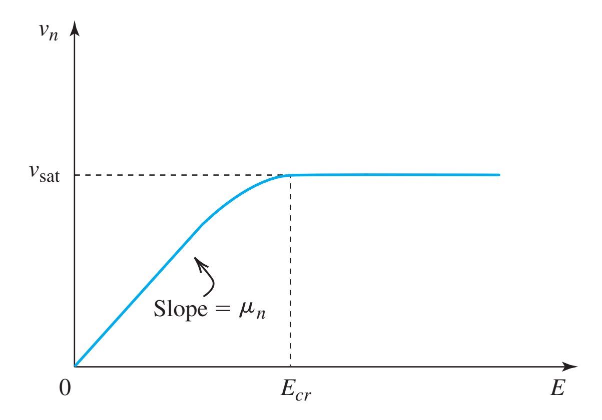


Figure 15.2 The velocity of electrons in the channel of an NMOS transistor reaches a constant value $v_{\text{sat}} \simeq 10^7$ cm/s when the electric field E reaches a critical value E_{cr}. A similar situation occurs for *p*-channel devices.

a linear to a constant ν is gradual, we shall assume for simplicity that ν saturates abruptly at $E = E_{cr}$.

The electric field E in a short-channel MOSFET can easily exceed E_{cr} even though V_{DD} is low. If we denote the value of v_{DS} at which velocity saturation occurs by $V_{DS\text{sat}}$, then from Eq. (15.2),

$$E_{cr} = \frac{V_{DS\text{sat}}}{L} \tag{15.3}$$

which when substituted in Eq. (15.1) provides

$$\nu_{\text{sat}} = \mu_n \left(\frac{V_{DS\text{sat}}}{L} \right) \tag{15.4}$$

or alternatively,

$$V_{DS\text{sat}} = \left(\frac{L}{\mu_n} \right) \nu_{\text{sat}} \tag{15.5}$$

Thus, $V_{DS\text{sat}}$ is a device parameter.

EXERCISE

15.3 Find $V_{DS\text{sat}}$ for an NMOS transistor fabricated in a 0.25-μm CMOS process with $\mu_n = 400\ \text{cm}^2/\text{V} \cdot \text{s}$. Let $L = 0.25\ \mu\text{m}$ and assume $\nu_{\text{sat}} = 10^7\ \text{cm/s}$.
Ans. 0.63 V

The i_D–v_{DS} Characteristics The i_D–v_{DS} equations of the MOSFET can be modified to include velocity saturation as follows. Consider a long-channel NMOS transistor operating in the triode region with v_{GS} set to a constant value V_{GS}. The drain current will be

$$i_D = \mu_n C_{ox} \left(\frac{W}{L} \right) v_{DS} \left[(V_{GS} - V_t) - \frac{1}{2} v_{DS} \right] \tag{15.6}$$

where we have for the time being neglected channel-length modulation. We know from our study in Section 5.1 that i_D will saturate at

$$v_{DS} = V_{OV} = V_{GS} - V_t \tag{15.7}$$

and the saturation current will be

$$i_D = \frac{1}{2} \mu_n C_{ox} \left(\frac{W}{L} \right) (V_{GS} - V_t)^2 \tag{15.8}$$

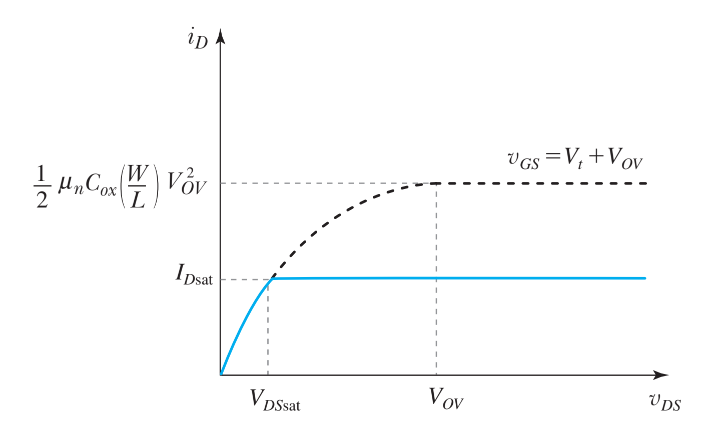


Figure 15.3 Velocity saturation causes the i_D–v_{DS} characteristic to saturate at V_{DSsat}. This early saturation results in a current I_{Dsat} that is lower than the value for a long-channel device.

This will also be the case in a short-channel device as long as the value of v_{DS} in Eq. (15.7) is lower than V_{DSsat}. That is, as long as

$$V_{OV} < V_{DSsat}$$

the current i_D will be given by Eqs. (15.6) and (15.8). If, on the other hand,

$$V_{OV} > V_{DSsat}$$

then velocity saturation kicks in at $v_{DS} = V_{DSsat}$ and i_D saturates at a value I_{Dsat}, as shown in Fig. 15.3. The value of I_{Dsat} can be obtained by substituting $v_{DS} = V_{DSsat}$ in Eq. (15.6),

$$I_{Dsat} = \mu_n C_{ox}\left(\frac{W}{L}\right)V_{DSsat}\left(V_{GS} - V_t - \frac{1}{2}V_{DSsat}\right) \tag{15.9}$$

This expression can be simplified by utilizing Eq. (15.5) to obtain

$$I_{Dsat} = WC_{ox}v_{sat}\left(V_{GS} - V_t - \frac{1}{2}V_{DSsat}\right) \tag{15.10}$$

Replacing V_{GS} in Eq. (15.9) with v_{GS}, and incorporating the channel-length modulation factor $(1 + \lambda v_{DS})$, we obtain a general expression for the drain current of an NMOS transistor operating in velocity saturation,

$$i_D = \mu_n C_{ox}\left(\frac{W}{L}\right)V_{DSsat}\left(v_{GS} - V_t - \frac{1}{2}V_{DSsat}\right)(1 + \lambda v_{DS}) \tag{15.11}$$

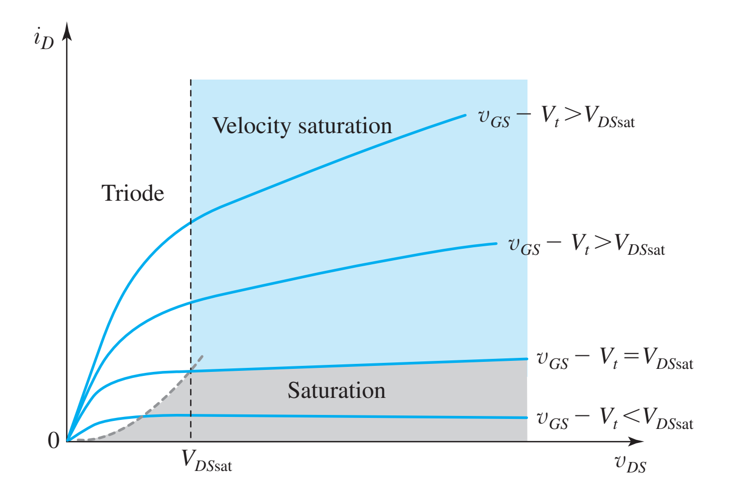


Figure 15.4 The i_D–v_{DS} characteristics of a short-channel MOSFET. Note the three different regions of operation: triode, saturation, and velocity saturation.

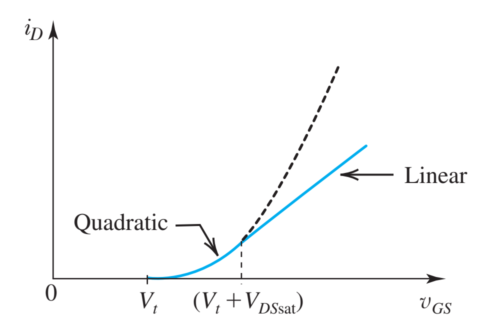


Figure 15.5 The i_D–v_{GS} characteristic of a short-channel NMOS transistor operating at $v_{DS} > V_{DSsat}$. Observe the quadratic and the linear portions of the characteristic. Also note that in the absence of velocity saturation, the quadratic curve would continue as shown with the broken line.

which applies for

$$v_{GS} - V_t \geq V_{DSsat} \qquad \text{and} \qquad v_{DS} \geq V_{DSsat} \tag{15.12}$$

Figure 15.4 shows a set of i_D–v_{DS} characteristic curves and clearly delineates the three regions of operation: triode, saturation, and velocity saturation.

Equation (15.11) indicates that in the velocity-saturation region, i_D is linearly related to v_{GS}. This is a major change from the quadratic relationship that characterizes operation in the saturation region. Figure 15.5 makes this point clearer by presenting a graph for i_D versus v_{GS} of a short-channel device operating at $v_{DS} > V_{DSsat}$. Observe that for $0 < v_{GS} - V_t \leq V_{DSsat}$, the MOSFET operates in the saturation region and i_D is related to v_{GS} by the familiar quadratic equation (Eq. 15.8). For $v_{GS} - V_t \geq V_{DSsat}$, the transistor enters the velocity-saturation region and i_D varies linearly with v_{GS} (Eq. 15.11).

Short-channel PMOS transistors undergo velocity saturation at the same value of ν_{sat} (approximately 10^7cm/s), but the effects on the device characteristics are less pronounced than in the NMOS case. This is due to the lower values of μ_p and the correspondingly higher values of E_{cr} and V_{DSsat}.

Example 15.1

Consider MOS transistors fabricated in a 0.25-μm CMOS process for which $V_{DD} = 2.5$ V, $V_{tn} = -V_{tp} = 0.5$ V, $\mu_n C_{ox} = 115\ \mu A/V^2$, $\mu_p C_{ox} = 30\ \mu A/V^2$, $\lambda_n = 0.06\ V^{-1}$, and $|\lambda_p| = 0.1\ V^{-1}$. Let $L = 0.25\ \mu m$ and $(W/L)_n = (W/L)_p = 1.5$. Measurements indicate that for the NMOS transistor, $V_{DSsat} = 0.63$ V, and for the PMOS device, $|V_{DSsat}| = 1$ V. Calculate the drain current obtained in each of the NMOS and PMOS transistors for $|V_{GS}| = |V_{DS}| = V_{DD}$. Compare with the values that would have been obtained in the absence of velocity saturation. Also give the range of v_{DS} for which i_D is saturated, with and without velocity saturation.

Solution

For the NMOS transistor, $V_{GS} = 2.5$ V results in $V_{GS} - V_{tn} = 2.5 - 0.5 = 2$ V, which is greater than V_{DSsat}. Also, $V_{DS} = 2.5$ V is greater than V_{DSsat}; thus both conditions in Eq. (15.12) are satisfied, and the NMOS transistor will be operating in the velocity-saturation region, and thus i_D is given by Eq. (15.11):

$$i_D = 115 \times 10^{-6} \times 1.5 \times 0.63 \times \left(2.5 - 0.5 - \frac{1}{2} \times 0.63\right) \times (1 + 0.06 \times 2.5) = 210.6\ \mu A$$

If velocity saturation were absent, the current would be

$$\begin{aligned} i_D &= \frac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)_n (v_{GS} - V_{tn})^2 (1 + \lambda v_{DS}) \\ &= \frac{1}{2} \times 115 \times 10^{-6} \times 1.5 \times (2.5 - 0.5)^2 \times (1 + 0.06 \times 2.5) \\ &= 396.8\ \mu A \end{aligned}$$

Thus, velocity saturation reduces the current level by nearly 50%! The saturation current, however, is obtained over a larger range of v_{DS}; specifically, for $v_{DS} = 0.63$ V to 2.5 V. (Of course, the current does not remain constant over this range because of channel-length modulation.) In the absence of velocity saturation, the current saturates at $V_{OV} = V_{GS} - V_t = 2$ V, and thus the saturation current is obtained over the range $v_{DS} = 2$ V to 2.5 V.

For the PMOS transistor, we see that since $|V_{GS}| - |V_t| = 2$ V and $|V_{DS}| = 2.5$ V are both larger than $|V_{DSsat}| = 1$ V the device will be operating in velocity saturation, and i_D can be obtained by adapting Eq. (15.11) as follows:

$$\begin{aligned} i_D &= (\mu_p C_{ox})\left(\frac{W}{L}\right)_p |V_{DSsat}|\left(|V_{GS}| - |V_{tp}| - \frac{1}{2}|V_{DSsat}|\right)(1 + |\lambda_p||V_{DS}|) \\ &= 30 \times 10^{-6} \times 1.5 \times 1 \times \left(2.5 - 0.5 - \frac{1}{2} \times 1\right)(1 + 0.1 \times 2.5) \\ &= 84.4\ \mu A \end{aligned}$$

Example 15.1 *continued*

Without velocity saturation, we have

$$i_D = \frac{1}{2}(\mu_p C_{ox})\left(\frac{W}{L}\right)_p (|V_{GS}| - |V_{tp}|)^2 (1 + |\lambda_p||V_{DS}|)$$

$$= \frac{1}{2} \times 30 \times 10^{-6} \times 1.5 \times (2.5 - 0.5)^2 (1 + 0.1 \times 2.5)$$

$$= 112.5\ \mu\text{A}$$

Thus velocity saturation reduces the current by 25% (which is less than in the case of the NMOS transistor), and the saturated current is obtained over the range $|V_{DS}| = 1$ V to 2.5 V. In the absence of velocity saturation, the saturated i_D would have been obtained for $|V_{DS}| = 2$ V to 2.5 V.

EXERCISE

15.4 Repeat the problem in Example 15.1 for transistors fabricated in a 0.13-μm CMOS process for which $V_{DD} = 1.2$ V, $V_{tn} = -V_{tp} = 0.4$ V, $\mu_n C_{ox} = 430\ \mu\text{A/V}^2$, $\mu_p C_{ox} = 110\ \mu\text{A/V}^2$, $\lambda_n = |\lambda_p| = 0.1\ \text{V}^{-1}$. Let $L = 0.13$ μm, $(W/L)_n = (W/L)_p = 1.5$, $V_{DS\text{sat}}$ (NMOS) = 0.34 V, and $V_{DS\text{sat}}$ (PMOS) = 0.6 V.
Ans. NMOS: $I_D = 154.7$ μA, compared to 231.2 μA without velocity saturation; saturation is obtained over the range $v_{DS} = 0.34$ V to 1.2 V, compared to $v_{DS} = 0.8$ V to 1.2 V in the absence of velocity saturation. PMOS: $I_D = 55.4$ μA compared to 59.1 μA, and $|v_{DS}| = 0.6$ V to 1.2 V compared to 0.8 V to 1.2 V.

Effect on the Inverter Characteristics The VTC of the CMOS inverter can be derived using the modified i_D–v_{DS} characteristics of the MOSFETs. The results, however, indicate relatively small changes from the VTC derived in Section 14.3 using the long-channel equations (see Rabaey et al., 2003, and Hodges et al., 2004), and we shall not pursue this subject here. The dynamic characteristics of the inverter, however, are significantly impacted by velocity saturation. This is because the current available to charge and discharge the equivalent load capacitance C is substantially reduced.

A Remark on the MOSFET Model The model derived above for short-channel MOSFETs is an approximate one, intended to enable the circuit designer to perform hand analysis to gain insight into circuit operation. Also, the model parameter values are usually obtained from measured data by means of a numerical curve-fitting process. As a result, the model applies only over a restricted range of terminal voltages.

Modeling short-channel MOSFETs is an advanced topic that is beyond the scope of this book. Suffice it to say that sophisticated models have been developed and are utilized by circuit simulation programs such as SPICE (see Appendix B). Circuit simulation is an essential step

in the design of integrated circuits. However, it is not a substitute for initial hand analysis and design.

15.1.4 Subthreshold Conduction

In our study of the NMOS transistor in Section 5.1, we assumed that current conduction between drain and source occurs only when v_{GS} exceeds V_t. That is, we assumed that for $v_{GS} < V_t$ no current flows between drain and source. This, however, turns out not to be the case, especially for deep-submicron devices. Specifically, for $v_{GS} < V_t$ a small current i_D flows. To be able to see this **subthreshold conduction**, we have redrawn the i_D–v_{GS} graph of Fig. 15.5, utilizing a logarithmic scale for i_D, as shown in Fig. 15.6. Observe that at low values of v_{GS}, the relationship between log i_D and v_{GS} is linear, indicating that i_D varies exponentially with v_{GS},

$$i_D = I_S e^{v_{GS}/nV_T} \tag{15.13}$$

where I_S is a constant, $V_T = kT/q$ is the thermal voltage $\simeq 25$ mV at room temperature, and n is a constant whose value falls in the range 1 to 2, depending on the material and structure of the device.[1]

Subthreshold conduction has been put to good use in the design of very-low-power circuits such as those needed for electronic watches. Generally speaking, however, subthreshold conduction is a problem in digital IC design. This is so for two reasons.

1. The nonzero current that flows for $v_{GS} = 0$ (see Fig. 15.6) causes the CMOS inverter to dissipate static power. To keep this **off current** as low as possible, V_t of the MOSFET is kept relatively high. This indeed is the reason why V_t has not been scaled by the same factor as that used for the channel length. Although the off current is low (10 pA to 100 pA) and the power dissipation per inverter is small, the problem becomes serious in chips with a billion transistors!

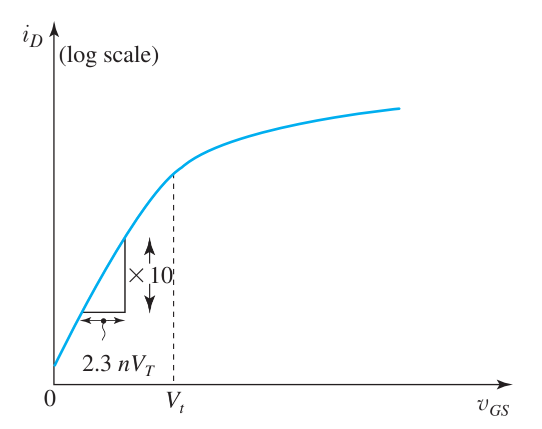


Figure 15.6 The i_D–v_{GS} characteristic of a short-channel MOSFET. To show the details of subthreshold conduction a logarithmic scale is needed for i_D.

[1]This relationship is reminiscent of the i_C–v_{BE} relationship of a BJT (Chapter 6). This is no coincidence, for the subthreshold conduction in a MOSFET is due to the lateral bipolar transistor formed by the source and drain diffusions with the substrate acting as the base region (see Fig. 5.1).

2. The nonzero current of a normally off transistor can cause the discharge of capacitors in dynamic MOS circuits. As we shall see in Section 15.5 and in the next chapter, dynamic logic and memory circuits rely on charge storage on capacitors for their proper operation. Thus, subthreshold conduction can disrupt the operation of such circuits.

EXERCISE

15.5 (a) Refer to Fig. 15.6 and to Eq. (15.13). Show that the inverse of the slope of the straight line representing subthreshold conduction is given by $2.3nV_T$ V per decade of current change.
(b) If measurements indicate $n = 1.22$ and $i_D = 100$ nA at $v_{GS} = 0.21$ V, find i_D at $v_{GS} = 0$.
(c) For a chip having 500 million transistors, find the current drawn from the 1.2-V supply V_{DD} as a result of subthreshold conduction. Hence estimate the resulting power dissipation.
Ans. (b) 0.1 nA; (c) 50 mA, 60 mW

15.1.5 Temperature, Voltage, and Process Variations

As we have seen in earlier chapters, temperature variations affect the i–v characteristics of a transistor. Besides affecting the thermal voltage V_T in subthreshold conduction, temperature variations impact the transistor threshold voltage V_t and the mobility μ. These effects, difficult to model in hand calculations, can have a significant impact on circuit performance, especially when a wide range of temperature environments is expected (we assume you would like your cell phone to work at the beach on a warm day, but also when your car breaks down on the way to a ski resort). Circuit simulators are usually equipped to model temperature effects, and designers should run their simulations at all extreme and expected temperatures.

Supply voltages also vary (e.g., to account for lower battery voltages during extended use), and it is common to require that circuits operate correctly for a range extending to *at least* $V_{DD} \pm 10\%$. Again, this condition can easily be simulated.

Finally, an increasingly important issue in CMOS design is that of process variations. Variations in threshold voltage, both at a small scale (transistor-to-transistor), medium scale (die-to-die), or large scale (wafer-to-wafer) should be expected.

15.1.6 Wiring: The Interconnect

The logic gates on a digital IC chip are connected together by metal wires[2] (see Appendix A). As well, the power-supply V_{DD} and ground are distributed throughout the chip by metal wires. Technology scaling into the deep-submicron range has caused these wires to behave not simply as wires! Specifically, the narrow wires typical of deep-submicron technologies exhibit nonzero resistance. The result is an *IR* drop on the V_{DD} line resulting in somewhat

[2]These are strips of metal deposited on an insulating surface on top of the chip. In modern digital ICs, as many as eight layers of such wiring are utilized.

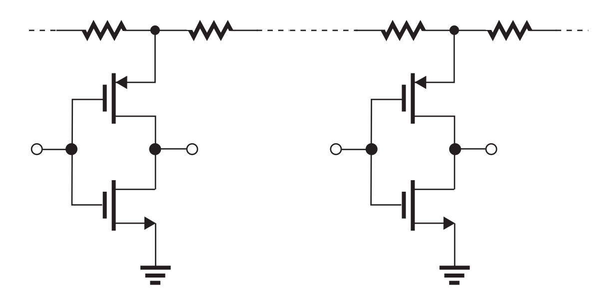


Figure 15.7 The power-supply line in a deep-submicron IC has nonzero resistance. The *IR* drops along the V_{DD} line cause the voltages delivered to various circuits to differ.

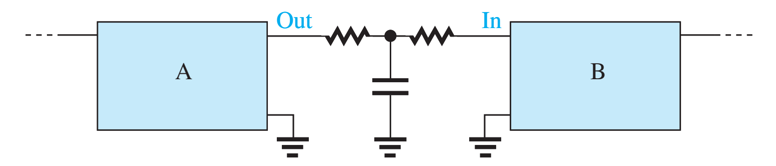


Figure 15.8 The interconnect (wire) between two circuit blocks, A and B, on an IC chip has finite resistance and a capacitance to ground.

different voltages being delivered to different parts of the chip, as shown in Fig. 15.7. This can have deleterious effects on the operation of the overall circuit.

Since chips fabricated in deep-submicron technologies can have hundreds of millions of gates, the wire connection between gates can be long. The resulting narrow and long **interconnect** lines have not only nonzero resistance but also capacitance to ground, as shown in Fig. 15.8. The resistance and capacitance of an interconnect line can cause a propagation delay approaching that of the logic gate itself. As well, the capacitance between adjacent wires can cause the signals on one wire to be coupled to the other, which can cause erroneous operation of logic circuits.

In short, the circuit designer of modern deep-submicron digital ICs has to concern herself not only with the logic-circuit design but also with the wiring or interconnect issues. Indeed, advanced textbooks on digital IC design devote entire chapters to this topic (see Rabaey et al., 2003, and Hodges et al., 2004). Our intent here is simply to point out that interconnect has become an important issue in digital IC design.

15.2 Digital IC Technologies, Logic-Circuit Families, and Design Methodologies

In our study of digital circuits, we have thus far concentrated on CMOS. This is reasonable in view of its dominance. Nevertheless, we will now take a broader view and survey other available digital circuit technologies. Not only will this help place CMOS in its proper context, it will also motivate the study, in the remainder of this chapter, of a number of other useful logic-circuit types. As well, we will briefly consider the methods digital IC designers employ to produce complex chips containing billions of transistors.

15.2.1 Digital IC Technologies and Logic-Circuit Families

The chart in Figure 15.9 shows the major IC technologies and logic-circuit families that are currently in use. The concept of a logic-circuit family perhaps needs a few words of explanation. The basic element of a logic-circuit family is the inverter. A family would include a variety of logic-circuit types made with the same technology, having a similar circuit structure, and exhibiting the same basic features. Each logic-circuit family offers a unique set of advantages and disadvantages. In the conventional style of designing systems, one selects an appropriate logic family (e.g., TTL, CMOS, or ECL) and attempts to implement as much of the system as possible using circuit modules (packages) that belong to this family. In this way, interconnection of the various packages is relatively straightforward. If, on the other hand, packages from more than one family are used, one has to design suitable *interface circuits*. The selection of a logic family is based on such considerations as logic flexibility, speed of operation, availability of complex functions, noise immunity, operating-temperature range, power dissipation, and cost. In the following, we make some brief remarks on each of the four technologies listed in the chart of Fig. 15.9.

CMOS Although shown as one of four possible technologies, this is not an indication of digital IC market share: CMOS technology is, by a very large margin, the most dominant of all the IC technologies available for digital circuit design. Although early microprocessors were made using NMOS logic (based on the inverter circuit in Fig. P14.28), CMOS has completely replaced NMOS. There are a number of reasons for this development, the most important of which is the much lower power dissipation of CMOS circuits. CMOS has also replaced bipolar as the technology of choice in digital system design and has made possible levels of integration (or circuit-packing densities) and a range of applications, neither of which would have been possible with bipolar technology. Furthermore, CMOS continues to advance, whereas there appear to be few innovations at the present time in bipolar digital circuits. Some of the reasons for CMOS displacing bipolar technology in digital applications are as follows.

1. CMOS logic circuits dissipate much less power than bipolar logic circuits and thus one can pack more CMOS circuits on a chip than is possible with bipolar circuits.
2. The high input impedance of the MOS transistor allows the designer to use charge storage as a means for the temporary storage of information in both logic and memory circuits. This technique cannot be used in bipolar circuits.

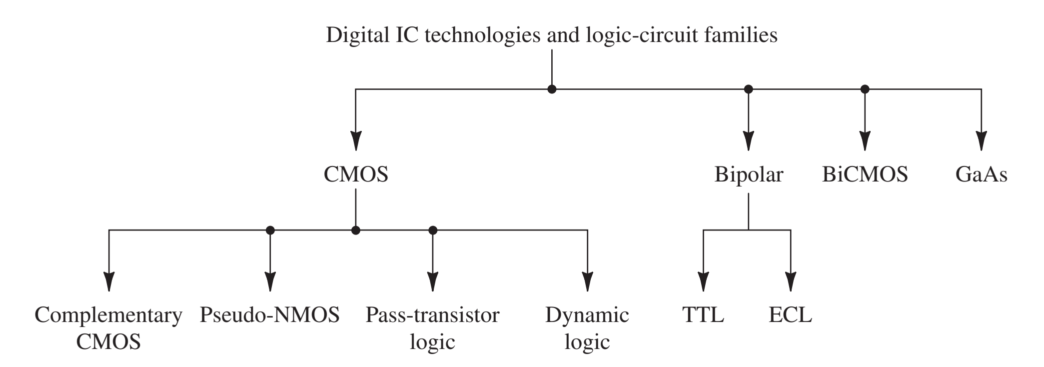


Figure 15.9 Digital IC technologies and logic-circuit families.

3. The feature size (i.e., minimum channel length) of the MOS transistor has decreased dramatically over the years, with some recently reported designs utilizing channel lengths as short as 22 nm. This permits very tight circuit packing and, correspondingly, very high levels of integration. A microprocessor chip reported in 2014 had 4.31 billion transistors.

Of the various forms of CMOS, complementary CMOS circuits, studied in Chapter 14, are the most widely used. They are available both as **small-scale-integrated (SSI)** circuit packages (containing 1–10 logic gates) and **medium-scale-integrated (MSI)** circuit packages (10–100 gates per chip) for assembling digital systems on printed-circuit boards. More significantly, complementary CMOS is used in **very-large-scale-integrated (VLSI)** logic (with millions of gates per chip) and memory-circuit design. In some applications, complementary CMOS is supplemented by one (or both) of two other MOS logic-circuit forms. These are pseudo-NMOS, so-named because of the similarity of its structure to NMOS logic, and pass-transistor logic, both of which will be studied in this chapter.

A fourth type of CMOS logic circuit utilizes dynamic techniques to obtain faster circuit operation, while keeping the power dissipation very low. Dynamic CMOS logic, which we shall study in Section 15.5, represents an area of growing importance. Lastly, CMOS technology is used in the design of memory chips, as will be detailed in Chapter 16.

Bipolar Two logic-circuit families based on the bipolar junction transistor are in some use at present: TTL and ECL. Transistor–transistor logic (TTL or T^2L) was for many years the most widely used logic-circuit family. Its decline was precipitated by the advent of the VLSI era. TTL manufacturers, however, fought back with the introduction of low-power and high-speed versions. In these relatively newer versions, the higher speeds of operation are made possible by preventing the BJT from saturating and thus avoiding the slow turnoff process of a saturated bipolar transistor. These nonsaturating versions of TTL utilize the Schottky diode discussed in Section 4.7.1 and are called Schottky TTL or variations of this name. Despite all these efforts, TTL is no longer a significant logic-circuit family and will not be studied in this book. However, the interested reader can find significant amounts of material on TTL on the book's website.

The other bipolar logic-circuit family in present use is emitter-coupled logic (ECL). It is based on the current-switch implementation of the inverter shown in Fig. 14.19. The basic element of ECL is the differential BJT pair studied in Chapter 9. Because ECL is basically a current-steering logic, and, correspondingly, also called **current-mode logic** (CML), in which saturation is avoided, very high speeds of operation are possible. Indeed, of all the commercially available logic-circuit families, ECL is the fastest. ECL is also used in VLSI circuit design when very high operating speeds are required and the designer is willing to accept higher power dissipation and increased silicon area. As such, ECL is considered an important specialty technology and will be discussed, albeit very briefly, in Section 15.6. Again, significant amounts of material on ECL can be found on the book's website.

BiCMOS BiCMOS combines the high operating speeds possible with BJTs (because of their inherently higher transconductance) with the low power dissipation and other excellent characteristics of CMOS. Like CMOS, BiCMOS allows for the implementation of both analog and digital circuits on the same chip. At present, BiCMOS is used to great advantage in special applications, where its high performance as a high-speed capacitive-current driver justifies the

more complex process technology it requires. A very brief discussion of BiCMOS is provided in Section 15.6, and more material can be found on the website.

Gallium Arsenide (GaAs) The high carrier mobility in GaAs results in very high speeds of operation. This has been demonstrated in a number of digital IC chips utilizing GaAs technology. It should be pointed out, however, that GaAs remains an "emerging technology," one that appears to have great potential but has not yet achieved such potential commercially. As such, it will not be studied in this book. Nevertheless, considerable material on GaAs devices and circuits, including digital circuits, can be found on the book's website.

THE INVISIBLE COMPUTER:

One may think that computer integrated circuits appear only in desktops, laptops, and mobile phones, but that is not the case! Virtually invisible to the casual observer is a vast number of computer chips called **microcontrollers**, which include a relatively high-speed processor (often of 8 bits, but increasingly 16 or 32), along with flash memory and flexible input/output circuitry. The input/output circuitry often includes A/D conversion. Microcontrollers operate within almost every modern appliance and computer peripheral: late-model automobiles include large numbers of networked microcontrollers for engine control, safety systems, stability, braking, and diagnostics. For example, in the 2012 Toyota Lexus there are about 100 such controllers.

15.2.2 Styles for Digital System Design

The conventional approach to designing digital systems consists of assembling the system using standard IC packages of various levels of complexity (and hence integration). Many systems have been built this way using, for example, TTL, SSI, and MSI packages. The advent of VLSI, in addition to providing the system designer with more powerful off-the-shelf components such as microprocessors and memory chips, has made possible alternative design styles. One such alternative is to opt for implementing part or all of the system using one or more *custom VLSI* chips. However, custom IC design is usually economically justified only when the production volume is large (greater than about 100,000 parts).

An intermediate approach, known as *semicustom design,* utilizes *gate-array* chips. These are integrated circuits containing 100,000 or more unconnected logic gates. Their interconnection can be achieved by a final metallization step (performed at the IC fabrication facility) according to a pattern specified by the user to implement the user's particular functional need. A more recently available type of gate array, known as a **field-programmable gate array (FPGA)**, can, as its name indicates, be programmed directly by the user. FPGAs provide a very convenient means for the digital system designer to implement complex logic functions in VLSI form without having to incur either the cost or the "turnaround time" inherent in custom and, to a lesser extent, in semicustom IC design.

15.2.3 Design Abstraction and Computer Aids

The design of very complex digital systems, whether on a single IC chip or using off-the-shelf components, is made possible by the use of different levels of design abstraction, and the

use of a variety of computer aids. To appreciate the concept of design abstraction, consider the process of designing a digital system using off-the-shelf packages of logic gates. The designer consults data sheets (in data books or on websites) to determine the input and output characteristics of the gates, their fan-in and fan-out limitations, and so on. In connecting the gates, the designer needs to adhere to a set of rules specified by the manufacturer in the data sheets. The designer does not need to consider, in a direct way, the circuit inside the gate package. In effect, the circuit has been abstracted in the form of a functional block that can be used as a component. This greatly simplifies system design. The digital IC designer follows a similar process. Circuit blocks are designed, characterized, and stored in a library as **standard cells**. These cells can then be used by the IC designer to assemble a larger subsystem (e.g., an adder or a multiplier), which in turn is characterized and stored as a functional block to be used in the design of an even larger system (e.g., an entire processor).

At every level of design abstraction, the need arises for simulation and other computer programs that help make the design process as automated as possible. Whereas SPICE is employed in circuit simulation, other software tools are utilized at other levels and in other phases of the design process. Although digital system design and design automation are outside the scope of this book, it is important that the reader appreciate the role of design abstraction and computer aids in digital design. They are what make it humanly possible to design a billion-transistor digital IC. Unfortunately, analog IC design does not lend itself to the same level of abstraction and automation. Each analog IC to a large extent has to be "handcrafted." As a result, the complexity and density of analog ICs remain much below what is possible in a digital IC.

Whatever approach or style is adopted in digital design, some familiarity with the various digital circuit technologies and design techniques is essential. Chapters 14, 15, and 16 aim to provide such a background.

15.3 Pseudo-NMOS Logic Circuits

15.3.1 The Pseudo-NMOS Inverter

Figure 15.10 shows a modified form of the CMOS inverter. Here, only Q_N is driven by the input voltage while the gate of Q_P is grounded, and Q_P acts as an active load for Q_N. Even before we examine the operation of this circuit in detail, an advantage over standard CMOS is obvious: Each input needs to be connected to the gate of only one transistor or, alternatively, only one additional transistor (an NMOS) will be needed for each additional gate input. Thus the area and delay penalties arising from increased fan-in in a standard CMOS will be reduced. This is indeed the motivation for exploring this modified inverter circuit. Note that we considered some aspects of this circuit earlier, in Example 14.3.

The inverter circuit of Fig. 15.10(a) resembles other forms of NMOS logic that consist of a driver transistor (Q_N) and a load transistor (in this case, Q_P); hence the name pseudo-NMOS. For comparison purposes, we shall briefly mention two older forms of NMOS logic. The earliest form, popular in the mid-1970s, utilized an enhancement MOSFET for the load element, in a topology whose basic inverter is shown in Fig. 15.10(b). This inverter circuit was the subject of Problem 14.28. It can be shown that its disadvantages include a relatively small logic swing, small noise margins, and high static power dissipation. For these reasons, this logic-circuit technology is virtually obsolete. It was replaced in the late 1970s and early 1980s with depletion-load NMOS circuits, in which a depletion NMOS transistor (see Section 5.4.5) with its gate connected to its source is used as the load element. The topology of the basic depletion-load inverter is shown in Fig. 15.10(c).

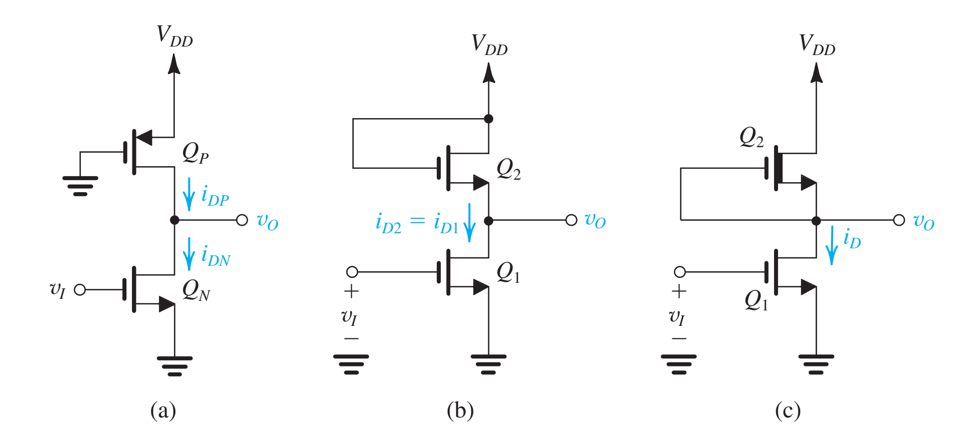


Figure 15.10 **(a)** The pseudo-NMOS logic inverter. **(b)** The enhancement-load (or saturated-load) NMOS inverter. **(c)** The depletion-load NMOS inverter.

It was initially expected that the depletion NMOS with $V_{GS} = 0$ would operate as a constant-current source and would thus provide an excellent load element.[3] However, it was quickly realized that the body effect in the depletion transistor causes its operation to deviate considerably from that of a constant-current source. Nevertheless, depletion-load NMOS circuits feature significant improvements over their enhancement-load counterparts, enough to justify the extra processing step required to fabricate the depletion devices (namely, ion-implanting the channel). Although depletion-load NMOS has been virtually replaced by CMOS, one can still see some depletion-load circuits in specialized applications. We will not study depletion-load NMOS logic here (the interested reader can refer to the website of this book).

The pseudo-NMOS inverter that we are about to study is similar to depletion-load NMOS, but with rather improved characteristics. It also has the advantage of being directly compatible with standard CMOS circuits.

15.3.2 Static Characteristics

The static characteristics of the pseudo-NMOS inverter can be derived in a manner similar to that used for standard CMOS. Toward that end, we note that the drain currents of Q_N and Q_P are given by

$$i_{DN} = \frac{1}{2}k_n(v_I - V_t)^2, \quad \text{for } v_O \geq v_I - V_t \quad \text{(saturation)} \tag{15.14}$$

$$i_{DN} = k_n\left[(v_I - V_t)v_O - \frac{1}{2}v_O^2\right] \quad \text{for } v_O \leq v_I - V_t \quad \text{(triode)} \tag{15.15}$$

[3] A constant-current load provides a capacitor-charging current that does not diminish as v_O rises toward V_{DD}, as is the case with a resistive load. Thus the value of t_{PLH} obtained with a current-source load is significantly lower than that obtained with a resistive load (see Problem 15.12). Of course, a resistive load, such as in the circuit studied in Example 14.2, is simply out of the question because of the very large silicon area it would occupy (equivalent to that of thousands of transistors!).

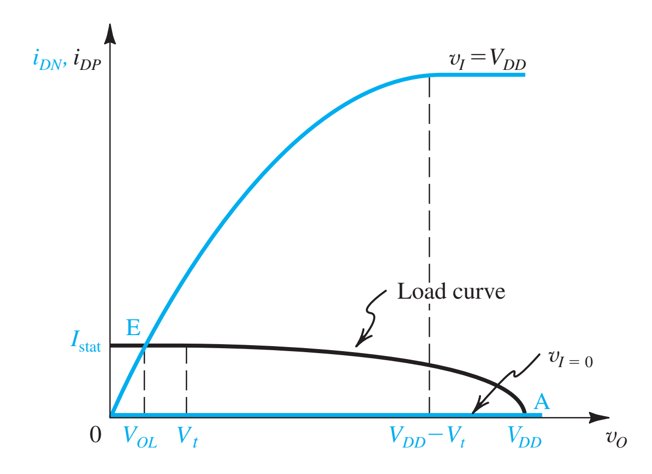


Figure 15.11 Graphical construction to determine the VTC of the inverter in Fig. 15.10(a).

$$i_{DP} = \frac{1}{2}k_p(V_{DD} - V_t)^2, \quad \text{for } v_O \le V_t \quad \text{(saturation)} \tag{15.16}$$

$$i_{DP} = k_p[(V_{DD} - V_t)(V_{DD} - v_O) - \frac{1}{2}(V_{DD} - v_O)^2], \quad \text{for } v_O \ge V_t \quad \text{(triode)} \tag{15.17}$$

where we have assumed that $V_{tn} = -V_{tp} = V_t$, and have used $k_n = k_n'(W/L)_n$ and $k_p = k_p'(W/L)_p$ to simplify matters.

To obtain the voltage-transfer characteristic of the inverter, we superimpose the load curve represented by Eqs. (15.16) and (15.17) on the i_D–v_{DS} characteristics of Q_N, which can be relabeled as i_{DN}–v_O and drawn for various values of $v_{GS} = v_I$. Such a graphical construction is shown in Fig. 15.11, where, to keep the diagram simple, we show the Q_N curves for only the two extreme values of v_I, namely, 0 and V_{DD}. Two observations follow:

1. The load curve represents a much lower saturation current (Eq. 15.16) than is represented by the corresponding curve for Q_N, namely, that for $v_I = V_{DD}$. This is a result of the fact that the pseudo-NMOS inverter is usually designed so that k_n is greater than k_p by a factor of 4 to 10. As we will show shortly, this inverter is of the so-called ratioed type,[4] and the ratio $r \equiv k_n/k_p$ determines all the breakpoints of the VTC, that is, V_{OL}, V_{IL}, V_{IH}, and so on, and thus determines the noise margins. Selection of a relatively high value for r reduces V_{OL} and widens the noise margins.
2. Although one tends to think of Q_P as acting as a constant-current source, it actually operates in saturation for only a small range of v_O, namely, $v_O \le V_t$. For the remainder of the v_O range, Q_P operates in the triode region.

Consider first the two extreme cases of v_I: When $v_I = 0$, Q_N is cut off and Q_P is operating in the triode region, though with zero current and zero drain–source voltage. Thus the operating point is that labeled A in Fig. 15.11, where $v_O = V_{OH} = V_{DD}$, the static current is zero, and the static power dissipation is zero. When $v_I = V_{DD}$, the inverter will operate at the point

[4]For the NMOS inverters such as that in Fig. 15.10(b), V_{OL} depends on the ratio of the transconductance parameters of the devices, that is, on the ratio $(k'(W/L))_{\text{driver}}/(k'(W/L))_{\text{load}}$. Such circuits are therefore known as *ratioed* logic circuits. Standard CMOS logic circuits do not have such a dependency and can therefore be called *ratioless*.

labeled E in Fig. 15.11. Observe that unlike standard CMOS, here V_{OL} is not zero, an obvious disadvantage. Another disadvantage is that the gate conducts current (I_{stat}) in the low-output state, and thus there will be static power dissipation ($P_D = I_{stat} \times V_{DD}$).

15.3.3 Derivation of the VTC

Figure 15.12 shows the VTC of the pseudo-NMOS inverter. As indicated, it has four distinct regions, labeled I through IV, corresponding to the different combinations of possible modes of operation of Q_N and Q_P. The four regions, the corresponding transistor modes of operation, and the conditions that define the regions are listed in Table 15.2. We shall utilize the information

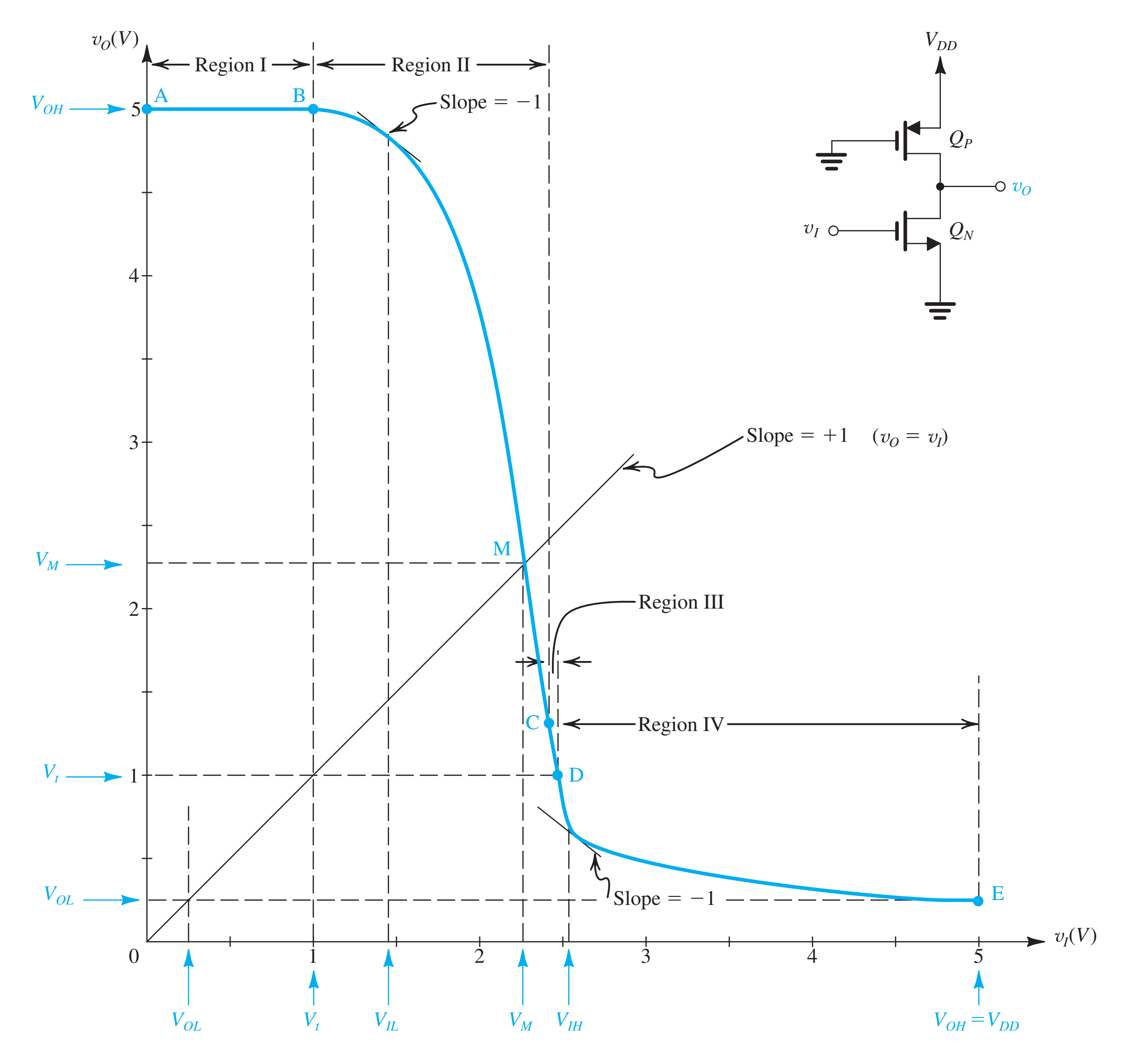


Figure 15.12 VTC for the pseudo-NMOS inverter. This curve is plotted for $V_{DD} = 5$ V, $V_{tn} = -V_{tp} = 1$ V, and $r = 9$.

Table 15.2 Regions of Operation of the Pseudo-NMOS Inverter

Region	Segment of VTC	Q_N	Q_P	Condition
I	AB	Cutoff	Triode	$v_I < V_t$
II	BC	Saturation	Triode	$v_O \geq v_I - V_t$
III	CD	Triode	Triode	$V_t \leq v_O \leq v_I - V_t$
IV	DE	Triode	Saturation	$v_O \leq V_t$

in this table together with the device equations given in Eqs. (15.14) through (15.17) to derive expressions for the various segments of the VTC and in particular for the important parameters that characterize the static operation of the inverter.

- **Region I (segment AB):**

$$v_O = V_{OH} = V_{DD} \tag{15.18}$$

- **Region II (segment BC):**

Equating i_{DN} from Eq. (15.14) and i_{DP} from Eq. (15.17) together with substituting $k_n = rk_p$, and with some manipulations, we obtain

$$v_O = V_t + \sqrt{(V_{DD} - V_t)^2 - r(v_I - V_t)^2} \tag{15.19}$$

The value of V_{IL} can be obtained by differentiating this equation and substituting $\partial v_O / \partial v_I = -1$ and $v_I = V_{IL}$:

$$V_{IL} = V_t + \frac{V_{DD} - V_t}{\sqrt{r(r+1)}} \tag{15.20}$$

The threshold voltage V_M is by definition the value of v_I for which $v_O = v_I$,

$$V_M = V_t + \frac{V_{DD} - V_t}{\sqrt{r+1}} \tag{15.21}$$

Finally, the end point of the region II segment (point C) can be found by substituting $v_O = v_I - V_t$ in Eq. (15.19), the condition for Q_N leaving saturation and entering the triode region.

- **Region III (segment CD)**

This is a short segment that is not of great interest. Point D is characterized by $v_O = V_t$.

- **Region IV (segment DE)**

Equating i_{DN} from Eq. (15.15) to i_{DP} from Eq. (15.16) and substituting $k_n = rk_p$ results in

$$v_O = (v_I - V_t) - \sqrt{(v_I - V_t)^2 - \frac{1}{r}(V_{DD} - V_t)^2} \tag{15.22}$$

The value of V_{IH} can be determined by differentiating this equation and setting $\partial v_O/\partial v_I = -1$ and $v_I = V_{IH}$,

$$V_{IH} = V_t + \frac{2}{\sqrt{3r}}(V_{DD} - V_t) \tag{15.23}$$

The value of V_{OL} can be found by substituting $v_I = V_{DD}$ into Eq. (15.22),

$$V_{OL} = (V_{DD} - V_t)\left[1 - \sqrt{1 - \frac{1}{r}}\right] \tag{15.24}$$

The static current conducted by the inverter in the low-output state is found from Eq. (15.16) as

$$I_{\text{stat}} = \frac{1}{2}k_p(V_{DD} - V_t)^2 \tag{15.25}$$

Finally, we can use Eqs. (15.20) and (15.24) to determine NM_L and Eqs. (15.18) and (15.23) to determine NM_H:

$$NM_L = V_t - (V_{DD} - V_t)\left[1 - \sqrt{1 - \frac{1}{r}} - \frac{1}{\sqrt{r(r+1)}}\right] \tag{15.26}$$

$$NM_H = (V_{DD} - V_t)\left(1 - \frac{2}{\sqrt{3r}}\right) \tag{15.27}$$

As a final observation, we note that since V_{DD} and V_t are determined by the process technology, the only design parameter for controlling the values of V_{OL} and the noise margins is the ratio r.

15.3.4 Dynamic Operation

Analysis of the inverter transient response to determine t_{PLH} with the inverter loaded by a capacitance C is identical to that of the complementary CMOS inverter. The capacitance will be charged by the current i_{DP}; we can determine an estimate for t_{PLH} by using the average value of i_{DP} over the range $v_O = 0$ to $v_O = V_{DD}/2$. The result is:

$$t_{PLH} = \frac{\alpha_p C}{k_p V_{DD}} \tag{15.28}$$

where

$$\alpha_p = 2\bigg/\left[\frac{7}{4} - 3\left(\frac{V_t}{V_{DD}}\right) + \left(\frac{V_t}{V_{DD}}\right)^2\right] \tag{15.29}$$

The case for the capacitor discharge is somewhat different because the current i_{DP} has to be subtracted from i_{DN} to determine the discharge current. The result is

$$t_{PHL} \simeq \frac{\alpha_n C}{k_n V_{DD}} \tag{15.30}$$

where

$$\alpha_n = 2 \Bigg/ \left[1 + \frac{3}{4}\left(1 - \frac{1}{r}\right) - \left(3 - \frac{1}{r}\right)\left(\frac{V_t}{V_{DD}}\right) + \left(\frac{V_t}{V_{DD}}\right)^2\right] \tag{15.31}$$

which, for a large value of r, reduces to

$$\alpha_n \simeq \alpha_p \tag{15.32}$$

Although these are similar formulas to those for the standard CMOS inverter, the pseudo-NMOS inverter has a special problem: Since k_p is r times smaller than k_n, t_{PLH} will be approximately r times larger than t_{PHL}. Thus the circuit exhibits an asymmetrical delay performance. Recall, however, that for gates with large fan-in, pseudo-NMOS requires fewer transistors and thus C can be smaller than in the corresponding standard CMOS gate.

15.3.5 Design

The design involves selecting the ratio r and the W/L for one of the transistors. The value of W/L for the other device can then be obtained using r. The design parameters of interest are V_{OL}, NM_L, NM_H, I_{stat}, P_D, t_{PLH}, and t_{PHL}. Important design considerations are as follows:

1. The ratio r determines all the breakpoints of the VTC; the larger the value of r, the lower V_{OL} is (Eq. 15.24) and the wider the noise margins are (Eqs. 15.26 and 15.27). However, a larger r increases the asymmetry in the dynamic response and, for a given $(W/L)_p$, makes the silicon area larger. Thus, selecting a value for r represents a compromise between noise margins on the one hand and silicon area and t_P on the other. Usually, r is selected in the range 4 to 10.
2. Once r has been determined, a value for $(W/L)_p$ or $(W/L)_n$ can be selected and the other determined. Here, one would select a small $(W/L)_n$ to keep the gate area small and thus obtain a small value for C. Similarly, a small $(W/L)_p$ keeps I_{stat} and P_D low. On the other hand, one would want to select larger W/L ratios to obtain low t_P and thus fast response. For usual (high-speed) applications, $(W/L)_p$ is selected so that I_{stat} is in the range of 50 μA to 100 μA, which for $V_{DD} = 1.8$ V results in P_D in the range of 90 μW to 180 μW.

15.3.6 Gate Circuits

Except for the load device, the pseudo-NMOS gate circuit is identical to the PDN of the complementary CMOS gate. Four-input, pseudo-NMOS NOR and NAND gates are shown in Fig. 15.13. Note that each requires five transistors compared to the eight used in standard CMOS. In pseudo-NMOS, NOR gates are preferred over NAND gates because the former do not utilize transistors in series and thus can be designed with minimum-size NMOS devices.

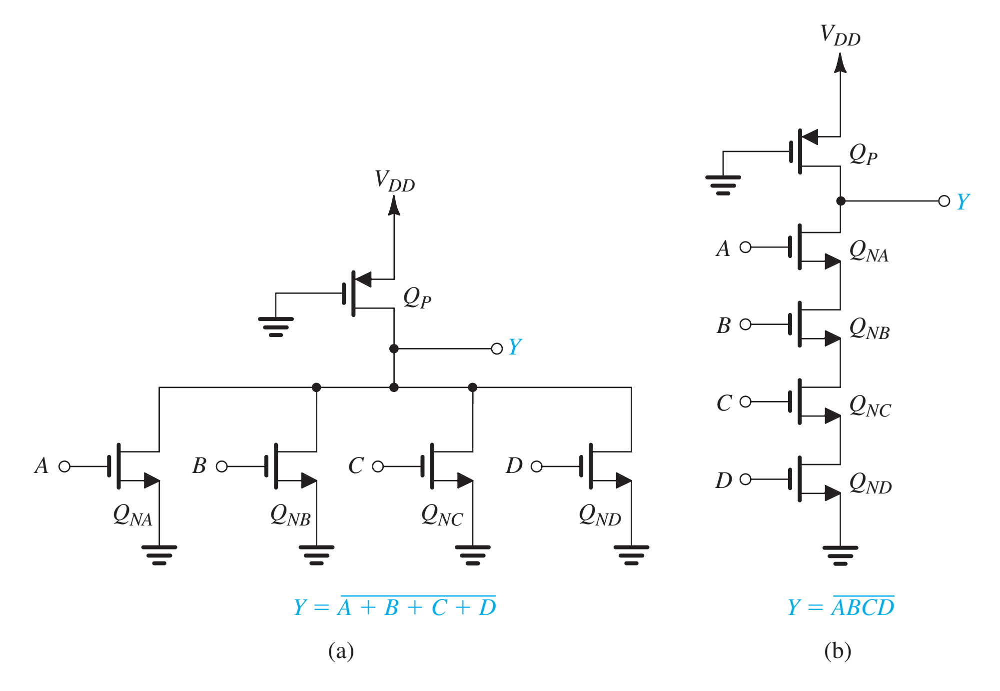


Figure 15.13 NOR and NAND gates of the pseudo-NMOS type.

15.3.7 Concluding Remarks

Pseudo-NMOS is particularly suited for applications in which the output remains high most of the time. In such applications, the static power dissipation can be reasonably low (since the gate dissipates static power only in the low-output state). Further, the output transitions that matter would presumably be high-to-low ones, where the propagation delay can be made as short as necessary. A particular application of this type can be found in the design of address decoders for memory chips (Section 16.4) and in read-only memories (Section 16.5).

Example 15.2

Consider a pseudo-NMOS inverter fabricated in a 0.25-μm CMOS technology for which $\mu_n C_{ox} = 115\ \mu A/V^2$, $\mu_p C_{ox} = 30\ \mu A/V^2$, $V_{tn} = -V_{tp} = 0.5$ V, and $V_{DD} = 2.5$ V. Let the W/L ratio of Q_P be (0.25 μm/0.25 μm) and $r = 9$. Find:

(a) V_{OH}, V_{OL}, V_{IL}, V_{IH}, V_M, NM_H, and NM_L
(b) $(W/L)_n$
(c) I_{stat} and P_D
(d) t_{PLH}, t_{PHL}, and t_P, assuming a total capacitance at the inverter output of 7 fF

Solution

(a) $V_{OH} = V_{DD} = 2.5$ V

V_{OL} is determined from Eq. (15.24) as

$$V_{OL} = (2.5 - 0.5)\left[1 - \sqrt{1 - \frac{1}{9}}\right] = 0.11 \text{ V}$$

V_{IL} is determined from Eq. (15.20) as

$$V_{IL} = 0.5 + \frac{2.5 - 0.5}{\sqrt{9(9+1)}} = 0.71 \text{ V}$$

V_{IH} is determined from Eq. (15.23) as

$$V_{IH} = 0.5 + \frac{2}{\sqrt{3 \times 9}} \times (2.5 - 0.5) = 1.27 \text{ V}$$

V_M is determined from Eq. (15.21) as

$$V_M = 0.5 + \frac{2.5 - 0.5}{\sqrt{9+1}} = 1.13 \text{ V}$$

The noise margins can now be determined as

$$NM_H = V_{OH} - V_{IH} = 2.5 - 1.27 = 1.23 \text{ V}$$

$$NM_L = V_{IL} - V_{OL} = 0.71 - 0.11 = 0.60 \text{ V}$$

Observe that the noise margins are not equal and that NM_L is rather low.

(b) The W/L ratio of Q_N can be found from

$$\frac{\mu_n C_{ox}(W/L)_n}{\mu_p C_{ox}(W/L)_p} = 9$$

$$\frac{115 \times (W/L)_n}{30 \times 1} = 9$$

Thus,

$$(W/L)_n = 2.35$$

(c) The dc current in the low-output state can be determined from Eq. (15.25) as

$$I_{\text{stat}} = \frac{1}{2} \times 30 \times 1(2.5 - 0.5)^2 = 60 \ \mu\text{A}$$

The static power dissipation can now be found from

$$P_D = I_{\text{stat}} V_{DD}$$
$$= 60 \times 2.5 = 150 \ \mu\text{W}$$

Example 15.2 *continued*

(d) The low-to-high propagation delay can be found by using Eqs. (15.28) and (15.29):

$$\alpha_p = 1.68$$

$$t_{PLH} = \frac{1.68 \times 7 \times 10^{-15}}{30 \times 10^{-6} \times 1 \times 2.5} = 0.16 \text{ ns}$$

The high-to-low propagation delay can be found by using Eqs. (15.30) and (15.31):

$$\alpha_n = 1.77$$

$$t_{PHL} = \frac{1.77 \times 7 \times 10^{-15}}{115 \times 10^{-6} \times 2.35 \times 2.5} = 0.02 \text{ ns}$$

Now, the propagation delay can be determined, as

$$t_P = \frac{1}{2}(0.16 + 0.02) = 0.09 \text{ ns}$$

Although the propagation delay is considerably greater than that of a standard CMOS inverter, this is not an entirely fair comparison: Recall that the advantage of pseudo-NMOS occurs in gates with large fan-in, not in a single inverter.

EXERCISES

15.6 While keeping r unchanged, redesign the inverter circuit of Example 15.2 to lower its static power dissipation to half the value found. Find the W/L ratios for the new design. Also find t_{PLH}, t_{PHL}, and t_P, assuming that C remains unchanged. Would the noise margins change?
Ans. $(W/L)_n = 1.18$; $(W/L)_p = 0.5$; 0.32 ns; 0.04 ns; 0.18 ns; no

15.7 Redesign the inverter of Example 15.2 using $r = 4$. Find V_{OL} and the noise margins. If $(W/L)_n = 0.375$ μm/0.25 μm, find $(W/L)_p$, I_{stat}, P_D, t_{PLH}, t_{PHL}, and t_P. Assume $C = 7$ fF.
Ans. $V_{OL} = 0.27$ V; $NM_L = 0.68$ V; $NM_H = 0.85$ V; $(W/L)_p = 1.44$; $I_{stat} = 86.3$ μA; $P_D = 0.22$ mW; $t_{PLH} = 0.11$ ns; $t_{PHL} = 0.03$ ns; $t_P = 0.07$ ns

15.4 Pass-Transistor Logic Circuits

A conceptually simple approach for implementing logic functions utilizes series and parallel combinations of switches that are controlled by input logic variables to connect the input and output nodes (see Fig. 15.14). Each of the switches can be implemented either by a

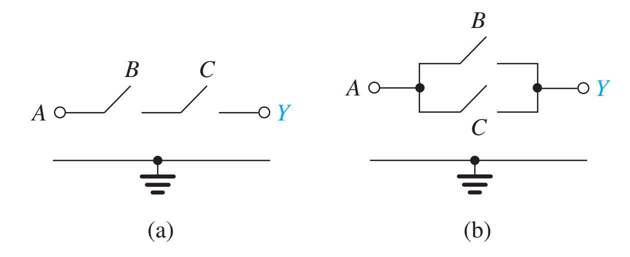


Figure 15.14 Conceptual pass-transistor logic gates. **(a)** Two switches, controlled by the input variables B and C, when connected in series in the path between the input node to which an input variable A is applied and the output node (with an implied load to ground) realize the function $Y = ABC$. **(b)** When the two switches are connected in parallel, the function realized is $Y = A(B + C)$.

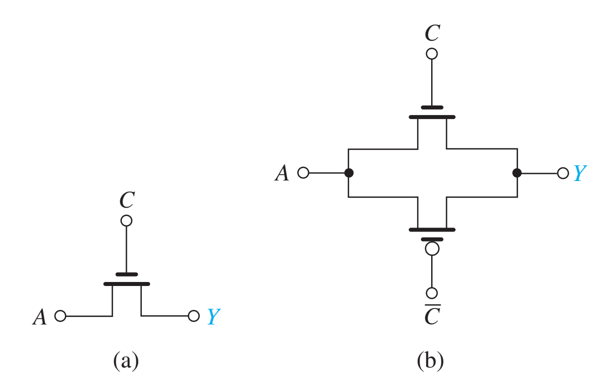


Figure 15.15 Two possible implementations of a voltage-controlled switch connecting nodes A and Y: **(a)** single NMOS transistor and **(b)** CMOS transmission gate.

single NMOS transistor [Fig. 15.15(a)] or by a pair of complementary MOS transistors connected in what is known as the **CMOS transmission-gate** configuration [Fig. 15.15(b)]. The result is a simple form of logic circuit that is particularly suited for some special logic functions and is frequently used in conjunction with standard CMOS logic to implement such functions efficiently: that is, with a lower total number of transistors than is possible with CMOS alone.

Because this form of logic utilizes MOS transistors in the series path from input to output, to *pass* or block signal transmission, it is known as *pass-transistor logic* (PTL). As mentioned earlier, CMOS transmission gates are frequently employed to implement the switches, giving this logic-circuit form the alternative name *transmission-gate logic*. The terms are used interchangeably independent of the actual implementation of the switches.

Though conceptually simple, pass-transistor logic circuits have to be designed with care. In the following, we shall study the basic principles of PTL circuit design and present examples of its application.

15.4.1 An Essential Design Requirement

An essential requirement in the design of PTL circuits is ensuring that *every circuit node has at all times a low-resistance path either to V_{DD} or to ground.* To appreciate this point, consider the situation depicted in Fig. 15.16(a): A switch S_1 (usually part of a larger PTL

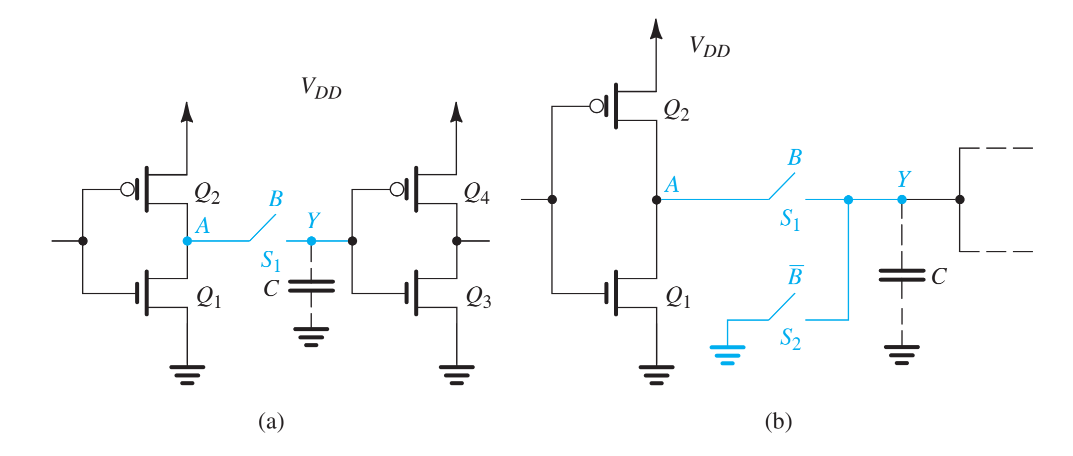


Figure 15.16 A basic design requirement of PTL circuits is that every node have, at all times, a low-resistance path to either ground or V_{DD}. Such a path does not exist in **(a)** when B is low and S_1 is open. It is provided in **(b)** through switch S_2.

network, not shown) is used to form the AND function of its controlling variable B and the variable A available at the output of a CMOS inverter. The output Y of the PTL circuit is shown connected to the input of another inverter. Now, if B is high, S_1 closes and $Y = A$. Node Y will then be connected either to V_{DD} (if A is high) through Q_2 or to ground (if A is low) through Q_1. But what happens when B goes low and S_1 opens? Node Y will now become a high-impedance node. If initially v_Y was zero, it will remain so. However, if initially v_Y was high at V_{DD}, this voltage will be maintained by the charge on the parasitic capacitance C, and Y will not be a logic 0 as required of the AND function $Y = AB$.

The problem can be easily solved by establishing for node Y a low-resistance path that is activated when B goes low, as shown in Fig. 15.16(b). Here, another switch, S_2, controlled by $\overline{B}$, is connected between Y and ground. When B goes low, S_2 closes and establishes a low-resistance path between Y and ground. The voltage v_Y will then be 0 volts, the proper output of the AND function when B is zero.

15.4.2 Operation with NMOS Transistors as Switches

Implementing the switches in a PTL circuit with single NMOS transistors results in a simple circuit with small area and small node capacitances. These advantages, however, are obtained at the expense of serious shortcomings in both the static characteristics and the dynamic performance of the resulting circuits. To illustrate, consider the circuit shown in Fig. 15.17, where an NMOS transistor Q is used to implement a switch connecting an input node with voltage v_I and an output node. The total capacitance between the output node and ground is represented by capacitor C. The switch is shown in the closed state with the control signal applied to its gate being high at V_{DD}. We wish to analyze the operation of the circuit as the input voltage v_I goes high (to V_{DD}) at time $t = 0$. We assume that initially the output voltage v_O is zero and capacitor C is fully discharged.[5]

[5]Although the MOS transistor is symmetric and its drain and source are interchangeable, it is always useful to know which terminal is functioning as the source and which as the drain. The terminal with the higher voltage in an NMOS transistor is the drain. The opposite is true for the PMOS transistor.

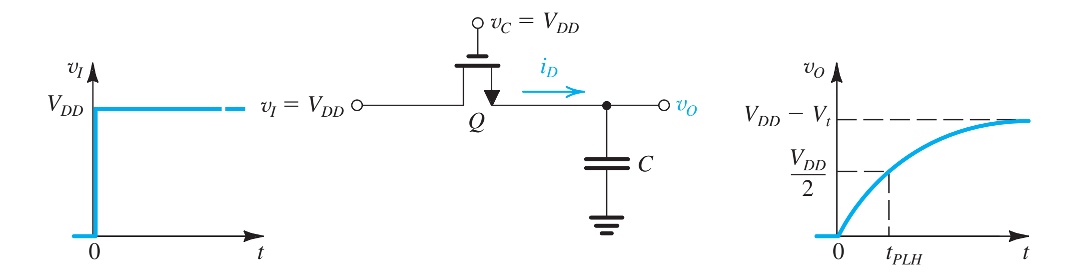


Figure 15.17 Operation of the NMOS transistor as a switch in the implementation of PTL circuits. This analysis is for the case with the switch closed (v_C is high) and the input going high ($v_I = V_{DD}$).

When v_I goes high, the transistor operates in the saturation mode and delivers a current i_D to charge the capacitor,

$$i_D = \frac{1}{2}k_n(V_{DD} - v_O - V_t)^2 \tag{15.33}$$

where $k_n = k_n'(W/L)$, and V_t is determined by the body effect since the source is at a voltage v_O relative to the body (which, though not shown, is connected to ground); thus (see Eq. 5.30),

$$V_t = V_{t0} + \gamma\left(\sqrt{v_O + 2\phi_f} - \sqrt{2\phi_f}\right) \tag{15.34}$$

Thus, initially (at $t = 0$), $V_t = V_{t0}$ and the current i_D is relatively large. However, as C charges up and v_O rises, V_t increases (Eq. 15.34) and i_D decreases. The latter effect is due to the increase both in v_O and in V_t. It follows that the process of charging the capacitor will be relatively slow. More seriously, observe from Eq. (15.33) that i_D reduces to zero when v_O reaches (V_{DD}–V_t). Thus the high output voltage (V_{OH}) will *not* be equal to V_{DD}; rather, it will be lower by V_t, and to make matters worse, the value of V_t can be as high as 1.5 to 2 times V_{t0}!

In addition to reducing the gate noise immunity, the low value of V_{OH} (commonly referred to as a "poor 1") has another detrimental effect: Consider what happens when the output node is connected to the input of a standard CMOS inverter (as was the case in Fig. 15.16). The low value of V_{OH} can cause Q_P of the load inverter to conduct. Thus the inverter will have a finite static current and static power dissipation.

The propagation delay t_{PLH} of the PTL gate of Fig. 15.17 can be determined as the time for v_O to reach $V_{DD}/2$. This can be calculated using techniques similar to those employed in the analysis of the CMOS inverter in Section 14.4, as will be illustrated shortly in an example.

Figure 15.18 shows the NMOS switch circuit when v_I is brought down to 0 V. We assume that initially $v_O = V_{DD}$. Thus at $t = 0+$, the transistor conducts and operates in the saturation region,

$$i_D = \frac{1}{2}k_n(V_{DD} - V_t)^2 \tag{15.35}$$

where we note that since the source is now at 0 V (note that the drain and source have interchanged roles), there will be no body effect, and V_t remains constant at V_{t0}. As C discharges, v_O decreases and the transistor enters the triode region at $v_O = V_{DD} - V_t$. Nevertheless, the capacitor discharge continues until C is fully discharged and $v_O = 0$. Thus, the NMOS transistor

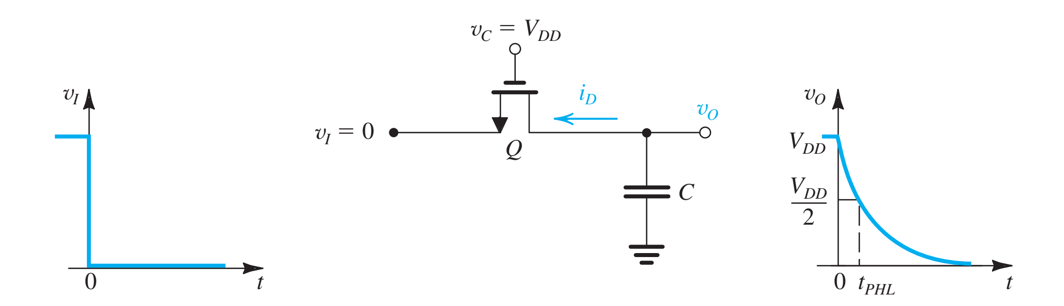


Figure 15.18 Operation of the NMOS switch as the input goes low ($v_I = 0$ V). Note that the drain of an NMOS transistor is always higher in voltage than the source; correspondingly, the drain and source terminals interchange roles in comparison to the circuit in Fig. 15.17.

provides $V_{OL} = 0$, or a "good 0." Again, the propagation delay t_{PHL} can be determined using usual techniques, as illustrated by the following example.

Example 15.3

Consider the NMOS transistor switch in the circuits of Figs. 15.17 and 15.18 to be fabricated in a technology for which $\mu_n C_{ox} = 50\ \mu\text{A/V}^2$, $\mu_p C_{ox} = 20\ \mu\text{A/V}^2$, $|V_{t0}| = 1$ V, $\gamma = 0.5\ \text{V}^{1/2}$, $2\phi_f = 0.6$ V, and $V_{DD} = 5$ V, where ϕ_f is a physical parameter. Let the transistor be of the minimum size for this technology, namely, 4 µm/2 µm, and assume that the total capacitance between the output node and ground is $C = 50$ fF.

(a) For the case with v_I high (Fig. 15.17), find V_{OH}.
(b) If the output feeds a CMOS inverter whose $(W/L)_p = 2.5\,(W/L)_n = 10$ µm/2 µm, find the static current of the inverter and its power dissipation when its input is at the value found in (a). Also find the inverter output voltage.
(c) Find t_{PLH}.
(d) For the case with v_I going low (Fig. 15.18), find t_{PHL}.
(e) Find t_P.

Solution

(a) Refer to Fig. 15.17. If V_{OH} is the value of v_O at which Q stops conducting,

$$V_{DD} - V_{OH} - V_t = 0$$

then,

$$V_{OH} = V_{DD} - V_t$$

where V_t is the value of the threshold voltage at a source–body reverse bias equal to V_{OH}. Using Eq. (15.34), we have

$$\begin{aligned} V_t &= V_{t0} + \gamma\left(\sqrt{V_{OH} + 2\phi_f} - \sqrt{2\phi_f}\right) \\ &= V_{t0} + \gamma\left(\sqrt{V_{DD} - V_t + 2\phi_f} - \sqrt{2\phi_f}\right) \end{aligned}$$

Substituting $V_{t0} = 1$, $\gamma = 0.5$, $V_{DD} = 5$, and $2\phi_f = 0.6$, we obtain a quadratic equation in V_t whose solution yields

$$V_t = 1.6 \text{ V}$$

Thus,

$$V_{OH} = 3.4 \text{ V}$$

Note that this represents a significant loss in signal amplitude.

(b) The load inverter will have an input signal of 3.4 V. Thus, its Q_P will conduct a current of

$$i_{DP} = \frac{1}{2} \times 20 \times \frac{10}{2}(5 - 3.4 - 1)^2 = 18 \ \mu\text{A}$$

where we have assumed Q_P to be operating in saturation, as we still expect v_O of the inverter to be close to 0 V. Thus, the static power dissipation of the inverter will be

$$P_D = V_{DD} i_{DP} = 5 \times 18 = 90 \ \mu\text{W}$$

The output voltage of the inverter can be found by noting that Q_N will be operating in the triode region. Equating its current to that of Q_P (i.e., 18 μA) enables us to determine the output voltage to be 0.08 V.

(c) To determine t_{PLH}, refer to Fig. 15.17. We need to find the current i_D at $t = 0$ (where $v_O = 0$, $V_t = V_{t0} = 1$ V) and at $t = t_{PLH}$ (where $v_O = 2.5$ V, V_t to be determined), as follows:

$$i_D(0) = \frac{1}{2} \times 50 \times \frac{4}{2} \times (5 - 1)^2 = 800 \ \mu\text{A}$$

$$V_t \ (\text{at } v_O = 2.5 \text{ V}) = 1 + 0.5\left(\sqrt{2.5 + 0.6} - \sqrt{0.6}\right) = 1.49 \text{ V}$$

$$i_D\left(t_{PLH}\right) = \frac{1}{2} \times 50 \times \frac{4}{2}(5 - 2.5 - 1.49)^2 = 50 \ \mu\text{A}$$

We can now compute the average discharge current as

$$i_D\big|_{\text{av}} = \frac{800 + 50}{2} = 425 \ \mu\text{A}$$

and t_{PLH} can be found as

$$t_{PLH} = \frac{C\left(V_{DD}/2\right)}{i_D\big|_{\text{av}}}$$

$$= \frac{50 \times 10^{-15} \times 2.5}{425 \times 10^{-6}} = 0.29 \text{ ns}$$

(d) Refer to the circuit in Fig. 15.18. Observe that, here, V_t remains constant at $V_{t0} = 1$ V. At $t = 0$, Q will be operating in saturation, and the drain current will be

$$i_D(0) = \frac{1}{2} \times 50 \times \frac{4}{2}(5 - 1)^2 = 800 \ \mu\text{A}$$

Example 15.3 *continued*

At $t = t_{PHL}$, Q will be operating in the triode region, and thus

$$i_D(t_{PHL}) = 50 \times \frac{4}{2}\left[(5-1) \times 2.5 - \frac{1}{2} \times 2.5^2\right]$$
$$= 688\ \mu A$$

Thus, the average discharge current is given by

$$i_D\big|_{av} = \frac{1}{2}(800 + 688) = 744\ \mu A$$

and t_{PHL} can be determined as

$$t_{PHL} = \frac{50 \times 10^{-15} \times 2.5}{744 \times 10^{-6}} = 0.17\ \text{ns}$$

(e) $t_P = \frac{1}{2}(t_{PLH} + t_{PHL}) = \frac{1}{2}(0.29 + 0.17) = 0.23$ ns

EXERCISE

15.8 Let the NMOS transistor switch in Fig. 15.17 be fabricated in a 0.18-μm CMOS process for which $V_{t0} = 0.5$ V, $\gamma = 0.3\ \text{V}^{1/2}$, $2\phi_f = 0.85$ V, and $V_{DD} = 1.8$ V. Find V_{OH}.
Ans. 1.15 V

15.4.3 Restoring the Value of V_{OH} to V_{DD}

Example 15.3 illustrates clearly the problem of signal-level loss and its deleterious effect on the operation of the succeeding CMOS inverter. Some rather ingenious techniques have been developed to restore the output level to V_{DD}. We shall briefly discuss two such techniques. One is circuit based and the other is based on process technology.

The circuit-based approach is illustrated in Fig. 15.19. Here, Q_1 is a pass-transistor controlled by input B. The output node of the PTL network is connected to the input of a standard CMOS inverter formed by Q_N and Q_P. A PMOS transistor Q_R, whose gate is controlled by the output voltage of the inverter, v_{O2}, has been added to the circuit. Observe that in the event that the output of the PTL gate, v_{O1}, is low (at ground), v_{O2} will be high (at V_{DD}), and Q_R will be off. On the other hand, if v_{O1} is high but not quite equal to V_{DD}, the output

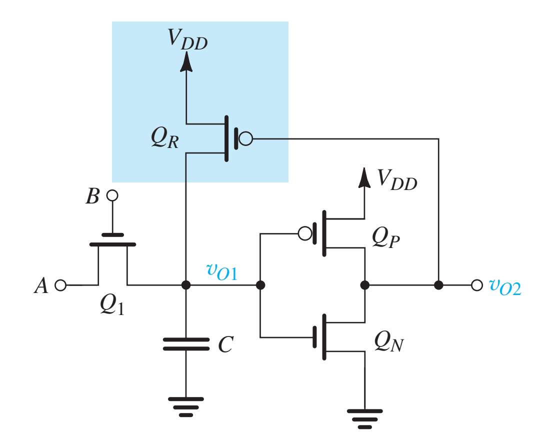


Figure 15.19 The use of transistor Q_R, connected in a feedback loop around the CMOS inverter, to restore the V_{OH} level, produced by Q_1, to V_{DD}.

of the inverter will be low (as it should be) and Q_R will turn on, supplying a current to charge C up to V_{DD}. This process will stop when $v_{O1} = V_{DD}$, that is, when the output voltage has been restored to its proper level. The "level-restoring" function performed by Q_R is frequently employed in MOS digital circuit design. It should be noted that although the description of operation is relatively straightforward, the addition of Q_R closes a "positive-feedback" loop around the CMOS inverter, and thus operation is more involved than it appears, especially during transients. Selection of a W/L ratio for Q_R is also a somewhat involved process, although normally k_r is selected to be much lower than k_n (say a third or a fifth as large). Intuitively, this is appealing, for it implies that Q_R will not play a major role in circuit operation, apart from restoring the level of V_{OH} to V_{DD}, as explained above. Transistor Q_R is said to be a "weak PMOS transistor." See Problem 15.35.

The other technique for correcting for the loss of the high-output signal level (V_{OH}) is a technology-based solution. Specifically, recall that the loss in the value of V_{OH} is equal to V_{tn}. It follows that we can reduce the loss by using a lower value of V_{tn} for the NMOS switches, and we can eliminate the loss altogether by using devices for which $V_{tn} = 0$. These **zero-threshold devices**, which can be fabricated by using ion implantation to control the value of V_{tn}, are known as **natural devices**. The problem of low-threshold devices, however, is the increased subthreshold conduction (Section 15.1.4) and the corresponding increase in static power dissipation.

15.4.4 The Use of CMOS Transmission Gates as Switches

Great improvements in static and dynamic performance are obtained when the switches are implemented with CMOS transmission gates. The transmission gate utilizes a pair of complementary transistors connected in parallel. It acts as an excellent switch, providing bidirectional current flow, and it exhibits an on-resistance that remains almost constant for wide ranges of input voltage. These characteristics make the transmission gate not only an excellent switch in digital applications but also an excellent analog switch in such applications as data converters and switched-capacitor filters (Chapter 17).

Before we analyze the transmission-gate circuit, it is useful to reflect on its origin. Recall that an NMOS transistor transmits the 0-V level to the output perfectly and thus produces a "good 0." It has difficulty, however, in passing the V_{DD} level, with the result that $V_{OH} = V_{DD} - V_t$

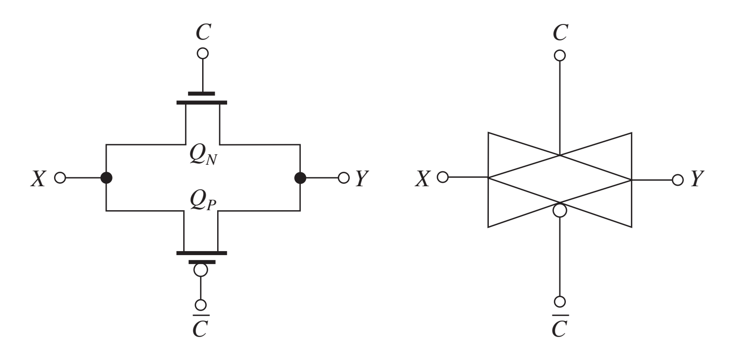


Figure 15.20 The CMOS transmission gate and its circuit symbol.

(a "poor 1"). It can be shown (see Problems 15.25 and 15.31) that a PMOS transistor does exactly the opposite; that is, it passes the V_{DD} level perfectly and thus produces a "good 1" but has trouble passing the 0-V level, thus producing a "poor 0." It is natural therefore to think that placing an NMOS and a PMOS transistor in parallel would produce good results in both the 0 and 1 cases.

Another way to describe the performance of the two transistor types is that the NMOS is good at pulling the output down to 0 V, while the PMOS is good at pulling the output up to V_{DD}. Interestingly, these are also the roles they play in the standard CMOS inverter.

Figure 15.20 shows the transmission gate together with its frequently used circuit symbol. The transmission gate is a bilateral switch that results in $v_Y = v_X$ when v_C is high (V_{DD}). In terms of logic variables, its function is described by

$$Y = X \quad \text{if} \quad C = 1$$

Figure 15.21(a) shows the transmission-gate switch in the "on" position with the input, v_I, rising to V_{DD} at $t = 0$. Assuming, as before, that initially the output voltage is zero, we see that Q_N will be operating in saturation and providing a charging current of

$$i_{DN} = \frac{1}{2}k_n(V_{DD} - v_O - V_{tn})^2 \tag{15.36}$$

where, as in the case of the single NMOS switch, V_{tn} is determined by the body effect,

$$V_{tn} = V_{t0} + \gamma\left(\sqrt{v_O + 2\phi_f} - \sqrt{2\phi_f}\right) \tag{15.37}$$

Transistor Q_N will conduct a diminishing current that reduces to zero at $v_O = V_{DD} - V_{tn}$. Observe, however, that Q_P operates with $V_{SG} = V_{DD}$ and is initially in saturation,

$$i_{DP} = \frac{1}{2}k_p\left(V_{DD} - |V_{tp}|\right)^2 \tag{15.38}$$

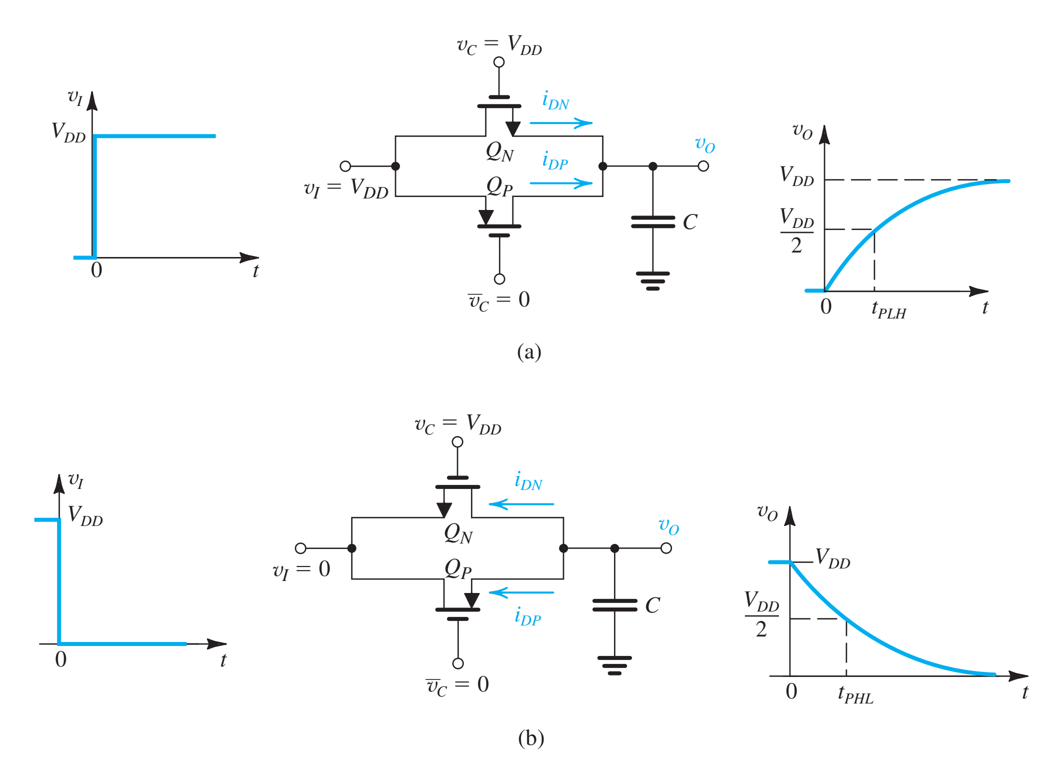


Figure 15.21 Operation of the transmission gate as a switch in PTL circuits with **(a)** v_I high and **(b)** v_I low.

where, since the body of Q_P is connected to V_{DD}, $|V_{tp}|$ remains constant at the value V_{t0}, assumed to be the same value as for the n-channel device. The total capacitor-charging current is the sum of i_{DN} and i_{DP}. Now, Q_P will enter the triode region at $v_O = |V_{tp}|$, but will continue to conduct until C is fully charged and $v_O = V_{OH} = V_{DD}$. Thus, the p-channel device will provide the gate with a "good 1." The value of t_{PLH} can be calculated using usual techniques, where we expect that as a result of the additional current available from the PMOS device, for the same value of C, t_{PLH} will be lower than in the case of the single NMOS switch. Note, however, that adding the PMOS transistor increases the value of C.

When v_I goes low, as shown in Fig. 15.21(b), Q_N and Q_P interchange roles. Analysis of the circuit in Fig. 15.21(b) will indicate that Q_P will cease conduction when v_O falls to $|V_{tp}|$, where $|V_{tp}|$ is given by

$$|V_{tp}| = V_{t0} + \gamma\left[\sqrt{V_{DD} - v_O + 2\phi_f} - \sqrt{2\phi_f}\right] \tag{15.39}$$

Transistor Q_N, however, continues to conduct until C is fully discharged and $v_O = V_{OL} = 0$ V, a "good 0."

We conclude that transmission gates provide far superior performance, both static and dynamic, than is possible with single NMOS switches. The price paid is increased circuit complexity, area, and capacitance.

EXERCISE

15.9 The transmission gate of Fig. 15.21(a) and 15.21(b) is fabricated in a CMOS process technology for which $k_n' = 50\ \mu\text{A/V}^2$, $k_p' = 20\ \mu\text{A/V}^2$, $V_{tn} = |V_{tp}|$, $V_{t0} = 1$ V, $\gamma = 0.5\ \text{V}^{1/2}$, $2\phi_f = 0.6$ V, and $V_{DD} = 5$ V. Let Q_N and Q_P be of the minimum size possible with this process technology, $(W/L)_n = (W/L)_p = 4\ \mu\text{m}/2\ \mu\text{m}$. The total capacitance at the output node is 70 fF. Utilize as many of the results of Example 15.3 as you need.

(a) What are the values of V_{OH} and V_{OL}?

(b) For the situation in Fig. 15.21(a), find $i_{DN}(0)$, $i_{DP}(0)$, $i_{DN}(t_{PLH})$, $i_{DP}(t_{PLH})$, and t_{PLH}.

(c) For the situation depicted in Fig. 15.21(b), find $i_{DN}(0)$, $i_{DP}(0)$, $i_{DN}(t_{PHL})$, $i_{DP}(t_{PHL})$, and t_{PHL}. At what value of v_O will Q_P turn off?

(d) Find t_P.

Ans. (a) 5 V, 0 V; (b) 800 μA, 320 μA, 50 μA, 275 μA, 0.24 ns; (c) 800 μA, 320 μA, 688 μA, 20 μA, 0.19 ns, 1.6 V; (d) 0.22 ns

Equivalent Resistance of the Transmission Gate Although the transmission gate is capable of passing the full 1 and 0 levels to the load capacitance, it is not a perfect switch. In particular, the transmission gate has a finite on-resistance. It is useful for us to obtain an estimate for this resistance. It can, for instance, be used together with the load capacitance as an alternative means to determining propagation delay. This approach is particularly useful in situations involving a network of inverters and transmission gates, as we shall shortly see.

To obtain an estimate of the resistance of the transmission gate, we shall consider the situation in Fig. 15.21(a), where the transmission gate is on and is passing a high input (V_{DD}) to the capacitor load. Transistor Q_N operates in saturation until the output voltage v_O reaches $(V_{DD} - V_{tn})$, at which time Q_N turns off; thus,

$$i_{DN} = \frac{1}{2}k_n(V_{DD} - V_{tn} - v_O)^2 \quad \text{for} \quad v_O \le V_{DD} - V_{tn} \tag{15.40}$$

$$i_{DN} = 0 \quad \text{for} \quad v_O \ge V_{DD} - V_{tn} \tag{15.41}$$

A gross estimate for the equivalent resistance of Q_N can be obtained by dividing the voltage across it, $(V_{DD} - v_O)$, by i_{DN}, and neglecting the body effect, that is, assuming V_{tn} remains constant; thus,

$$R_{Neq} = \frac{V_{DD} - v_O}{\frac{1}{2}k_n(V_{DD} - V_{tn} - v_O)^2} \quad \text{for} \quad v_O \le V_{DD} - V_{tn} \tag{15.42}$$

and

$$R_{Neq} = \infty \quad \text{for} \quad v_O \ge V_{DD} - V_{tn} \tag{15.43}$$

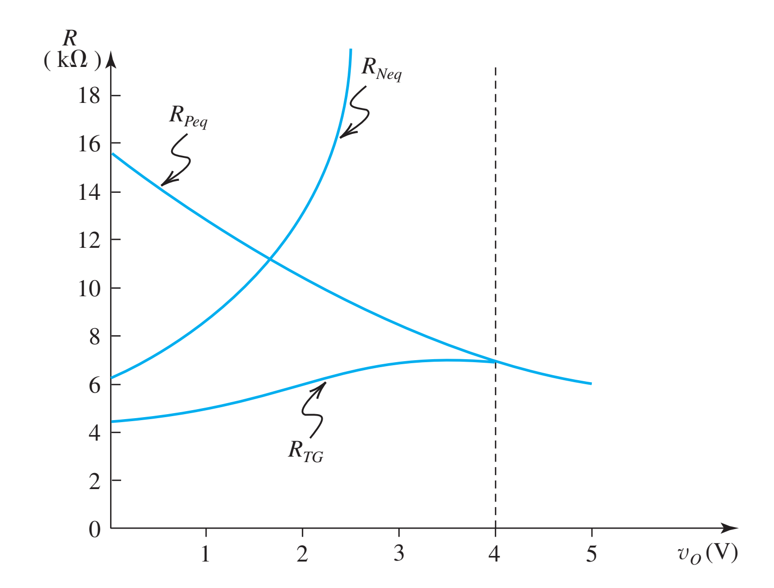


Figure 15.22 Plot of the equivalent resistances of the two transistors of the transmission gate in Fig. 15.21(a) and the overall resistance R_{TG} versus v_O. The data apply to the situation specified in Exercise 15.10.

Transistor Q_P will operate in saturation until $v_O = |V_{tp}|$, after which it enters the triode region; thus,

$$i_{DP} = \frac{1}{2}k_p(V_{DD} - |V_{tp}|)^2 \quad \text{for} \quad v_O \leq |V_{tp}| \tag{15.44}$$

$$i_{DP} = k_p\left[(V_{DD} - |V_{tp}|)(V_{DD} - v_O) - \frac{1}{2}(V_{DD} - v_O)^2\right] \quad \text{for} \quad v_O \geq |V_{tp}| \tag{15.45}$$

A gross estimate for the resistance of Q_P can be obtained by dividing the voltage across it, $(V_{DD} - v_O)$, by i_{DP}; thus,

$$R_{Peq} = \frac{V_{DD} - v_O}{\frac{1}{2}k_p(V_{DD} - |V_{tp}|)^2} \quad \text{for} \quad v_O \leq |V_{tp}| \tag{15.46}$$

$$R_{Peq} = \frac{1}{k_p\left[V_{DD} - |V_{tp}| - \frac{1}{2}(V_{DD} - v_O)\right]} \quad \text{for} \quad v_O \geq |V_{tp}| \tag{15.47}$$

Finally, the equivalent resistance R_{TG} of the transmission gate can be obtained as the parallel equivalent of R_{Neq} and R_{Peq},

$$R_{TG} = R_{Neq} \,\|\, R_{Peq} \tag{15.48}$$

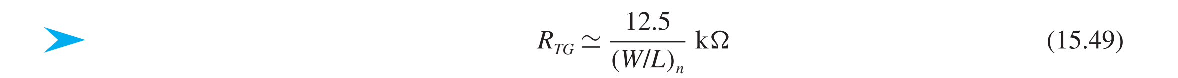
Obviously, R_{TG} is a function of the output voltage v_O. As an example, we show in Fig. 15.22 a plot for R_{TG} for the transmission gate analyzed in Exercise 15.9. Observe that R_{TG} remains relatively constant over the full range of v_O. The average value of R_{TG} over the range $v_O = 0$ to $V_{DD}/2$ can be used to determine t_{PLH}, as illustrated in Exercise 15.10.

EXERCISE

15.10 For the transmission gate analyzed in Exercise 15.9, whose equivalent resistance for capacitor charging is plotted in Fig. 15.22, use the average resistance value over the range $v_O = 0$ V to 2.5 V to determine t_{PLH}. Compare the result to that obtained using average currents in Exercise 15.9. Note that from the graph, $R_{TG} = 4.5$ kΩ at $v_O = 0$ V, and $R_{TG} = 6.5$ kΩ at $v_O = 2.5$ V. Recall that $t_{PLH} = 0.69RC$.
Ans. $t_{PLH} = 0.27$ ns, very close to the value of 0.24 ns obtained in Exercise 15.9

The expression for R_{TG} derived above applies only to the case of capacitor charging. A similar analysis can be performed for the case of capacitor discharge illustrated in Fig. 15.21(b). The resulting value of R_{TG} is close to that obtained above (see Problem 15.38).

Similar to the empirical formulas for R_N and R_P of the CMOS inverter (Eqs. 14.56 and 14.57), there is a simple empirical formula for R_{TG} that applies for both capacitor charging and discharging and for all modern submicron technologies (see Hodges et al., 2004), namely,



$$R_{TG} \simeq \frac{12.5}{(W/L)_n} \text{ k}\Omega \qquad (15.49)$$

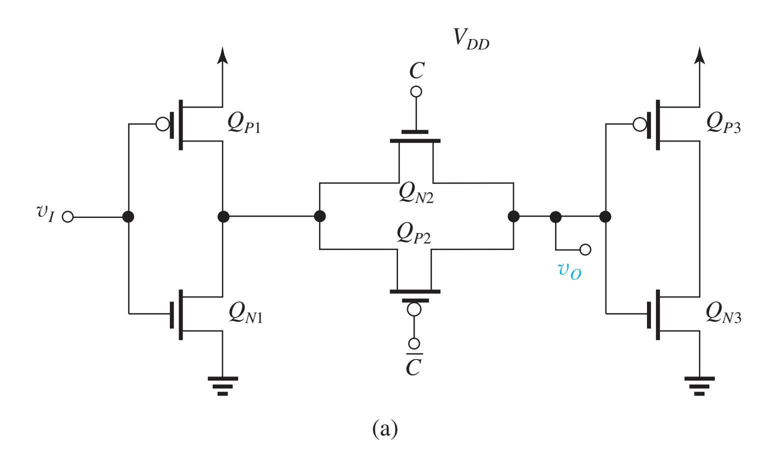
EXERCISE

15.11 Use Eq. (15.49) to estimate the value of R_{TG} for a transmission gate fabricated in a 0.18-μm CMOS technology with $(W/L)_n = (W/L)_p = 1.5$.
Ans. 8.3 kΩ

Having an estimate of the resistance of the transmission gate enables us to calculate the propagation delay of a signal path containing one or more transmission gates. Figure 15.23(a) shows one such circuit. It consists of a transmission gate connecting the output of an inverter to the input of another. We are interested in finding the propagation delay from the input of the first inverter to the input of the second as we apply a negative-going step to the input of the first inverter.

Fig. 15.23(b) shows the equivalent circuit where R_{P1} is the equivalent resistance of Q_{P1}, R_{TG} is the equivalent resistance of the transmission gate, C_{out1} is the output capacitance of the driver inverter, C_{TG1} and C_{TG2} are the capacitances introduced by the transmission gate at its input and output, respectively, and C_{in2} is the input capacitance of the load inverter. Observe that the circuit takes the form of an RC ladder network. A simple formula has been developed for calculating the delay of an arbitrarily long RC ladder network such as that shown in Fig. 15.24 having three sections. Known as the **Elmore delay formula**, it gives for the ladder in Fig. 15.24

$$t_P = 0.69[C_1R_1 + C_2(R_1 + R_2) + C_3(R_1 + R_2 + R_3)] \qquad (15.50)$$



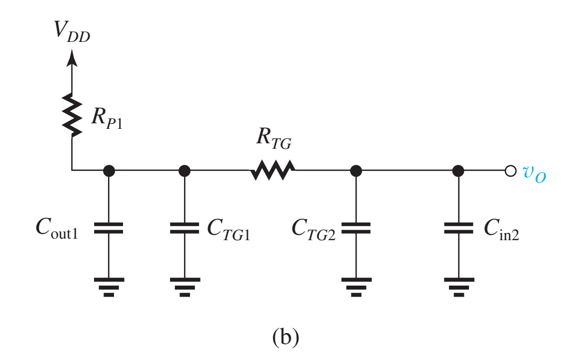


Figure 15.23 **(a)** A transmission gate connects the output of a CMOS inverter to the input of another. **(b)** Equivalent circuit for the purpose of analyzing the propagation delay of the circuit in **(a)**.

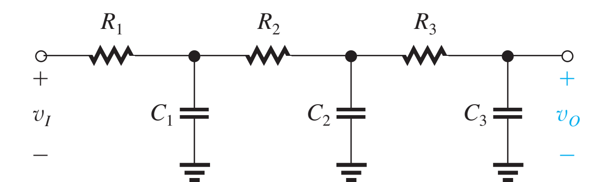


Figure 15.24 A three-section *RC* ladder network.

Applying the Elmore formula to the two-stage ladder in Fig. 15.23(b) gives

$$t_P = 0.69[(C_{\text{out1}} + C_{TG1})R_{P1} + (C_{\text{in2}} + C_{TG2})(R_{P1} + R_{TG})] \tag{15.51}$$

Finally, the factor 0.69 is usually dropped on the assumption that the input is a ramp rather than a step voltage.

EXERCISE

15.12 The circuit in Fig. 15.23 is fabricated in a 0.13-μm CMOS technology; Q_P of the first inverter has $W/L = 2$, and both transistors of the transmission gate have $W/L = 1$. The capacitances have been estimated to be $C_{out1} = 10$ fF, $C_{TG1} = C_{TG2} = 5$ fF, and $C_{in2} = 10$ fF. Use the empirical formulas to obtain the values of R_{P1} and R_{TG}. Then, determine an estimate for t_P assuming a ramp input.
Ans. $R_{P1} = 15$ kΩ; $R_{TG} = 12.5$ kΩ; $t_P = 0.64$ ns

15.4.5 Examples of Pass-Transistor Logic Circuits

We conclude this section by showing examples of PTL logic circuits. Figure 15.25 shows a PTL realization of a two-to-one multiplexer: Depending on the logic value of C, either A or B is connected to the output Y. The circuit realizes the Boolean function

$$Y = CA + \overline{C}B$$

Our second example is an efficient realization of the exclusive-OR (XOR) function. The circuit, shown in Fig. 15.26, utilizes four transistors in the transmission gates and another four for the two inverters needed to generate the complements $\overline{A}$ and $\overline{B}$, for a total of eight transistors. Note that 12 transistors are needed in the realization with standard CMOS.

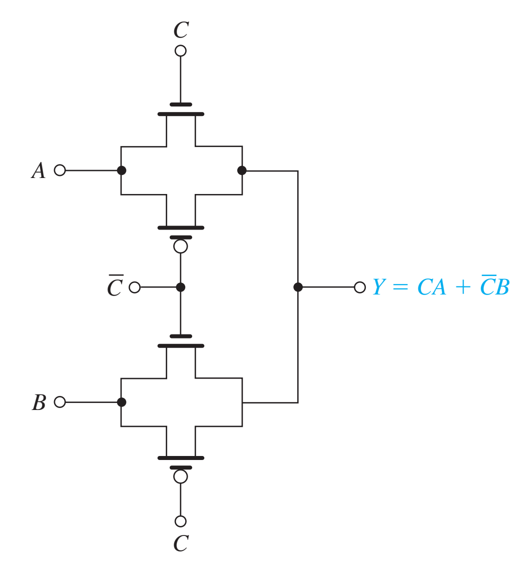


Figure 15.25 Realization of a two-to-one multiplexer using pass-transistor logic.

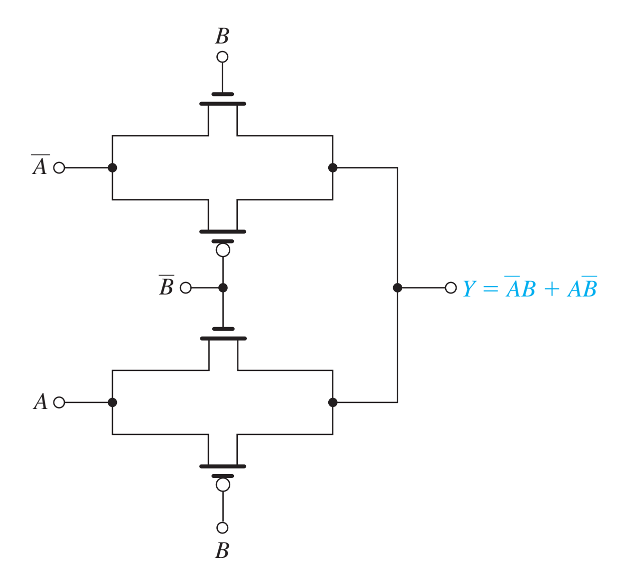


Figure 15.26 Realization of the XOR function using pass-transistor logic.

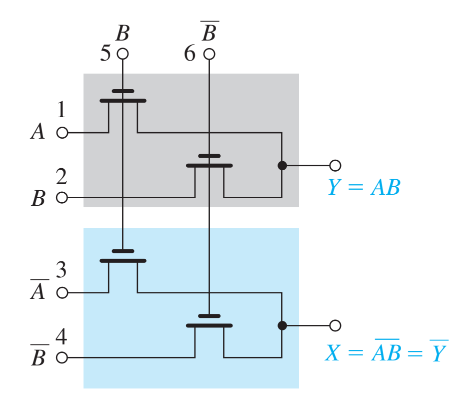


Figure 15.27 An example of a pass-transistor logic gate utilizing both the input variables and their complements. This type of circuit is therefore known as complementary pass-transistor logic, or CPL. Note that both the output function and its complement are generated.

Our final PTL example is the circuit shown in Fig. 15.27. It uses NMOS switches with low or zero threshold. Observe that both the input variables and their complements are employed and that the circuit generates both the Boolean function and its complement. Thus this form of circuit is known as **complementary pass-transistor logic (CPL)**. The circuit consists of two identical networks of pass transistors with the corresponding transistor gates controlled by the same signal (B and $\overline{B}$). The inputs to the PTL, however, are complemented: A and B for the first network, and $\overline{A}$ and $\overline{B}$ for the second. The circuit shown realizes both the AND and NAND functions.

EXERCISE

15.13 Consider the circuit in Fig. 15.27, and for each case, find Y and $\overline{Y}$. The input signals are changed as follows:

(a) The signals at terminals 5 and 6 are interchanged ($\overline{B}$ applied to 5 and B applied to 6). All the rest are the same.

(b) The signals at terminals 5 or 6 are interchanged as in (a), and the signals at 2 and 4 are changed to $\overline{A}$ and A, respectively. All the rest remain the same.

Ans. (a) $Y = A + B, \overline{Y} = \overline{A}\,\overline{B} = \overline{A + B}$ (i.e., OR–NOR); (b) $Y = A\overline{B} + \overline{A}B, \overline{Y} = \overline{A}\,\overline{B} + AB$ (i.e., XOR–XNOR)

15.4.6 A Final Remark

Although the use of zero-threshold devices solves the problem of the loss of signal levels when NMOS switches are used, the resulting circuits can be much more sensitive to noise and other effects, such as leakage currents resulting from subthreshold conduction.



15.5 Dynamic MOS Logic Circuits

The logic circuits that we have studied thus far are of the static type. In a static logic circuit, every node has, at all times, a low-resistance path to V_{DD} or ground. By the same token, the voltage of each node is well defined at all times, and no node is left floating. Static circuits do not need clocks (i.e., periodic timing signals) for their operation, although clocks may be present for other purposes. In contrast, the dynamic logic circuits we are about to discuss rely on the storage of signal voltages on parasitic capacitances at certain circuit nodes. Since charge will leak away with time, the circuits need to be *periodically refreshed;* thus the presence of a clock with a certain specified minimum frequency is essential.

To place dynamic logic-circuit techniques into perspective, let's take stock of the various styles we have studied for logic circuits. Standard CMOS excels in nearly every performance category: It is easy to design, has the maximum possible logic swing, is robust from a noise-immunity standpoint, dissipates no static power, and can be designed to provide equal low-to-high and high-to-low propagation delays. Its main disadvantage is the requirement of two transistors for each additional gate input, which for high fan-in gates can make the chip area large and increase the total capacitance and, correspondingly, the propagation delay and the dynamic power dissipation. Pseudo-NMOS reduces the number of required transistors at the expense of static power dissipation. Pass-transistor logic can result in simple small-area circuits but is limited to special applications and requires the use of CMOS inverters to restore signal levels, especially when the switches are simple NMOS transistors. The dynamic logic techniques studied in this section maintain the low device count of pseudo-NMOS while reducing the static power dissipation to zero. As will be seen, this is achieved at the expense of more complex, and less robust, design.

15.5.1 The Basic Principle

Figure 15.28(a) shows the basic dynamic logic gate. It consists of a pull-down network (PDN) that realizes the logic function in exactly the same way as the PDN of a standard CMOS gate or a pseudo-NMOS gate. Here, however, we have two switches in series that are periodically operated by the clock signal ϕ whose waveform is shown in Fig. 15.28(b). When ϕ is low, Q_p is turned on, and the circuit is said to be in the setup or **precharge phase**. When ϕ is high, Q_p is off and Q_e turns on, and the circuit is in the **evaluation phase**. Finally, note that C_L denotes the total capacitance between the output node and ground.

During precharge, Q_p conducts and charges capacitance C_L so that at the end of the precharge interval, the voltage at Y is equal to V_{DD}. Also during precharge, the inputs A, B, and C are allowed to change and settle to their proper values. Observe that because Q_e is off, no path to ground exists.

During the evaluation phase, Q_p is off and Q_e is turned on. Now, if the input combination is one that corresponds to a high output, the PDN does not conduct (just as in a standard CMOS gate) and the output remains high at V_{DD}; thus $V_{OH} = V_{DD}$. Observe that no low-to-high propagation delay is required, thus $t_{PLH} = 0$. On the other hand, if the combination of inputs is one that corresponds to a low output, the appropriate NMOS transistors in the PDN will conduct and establish a path between the output node and ground through the on transistor Q_e. Thus C_L will be discharged through the PDN, and the voltage at the output node will reduce to $V_{OL} = 0$ V. The high-to-low propagation delay t_{PHL} can be calculated in exactly the same way as for a standard CMOS circuit, except that here we have an additional transistor, Q_e, in the series path to ground. Although this will increase the delay slightly, the increase will be more than offset by the reduced capacitance at the output node as a result of the absence of the PUN.

As an example, we show in Fig. 15.28(c) the circuit that realizes the function $Y = \overline{A + BC}$. Sizing of the PDN transistors often follows the same procedure employed in the design

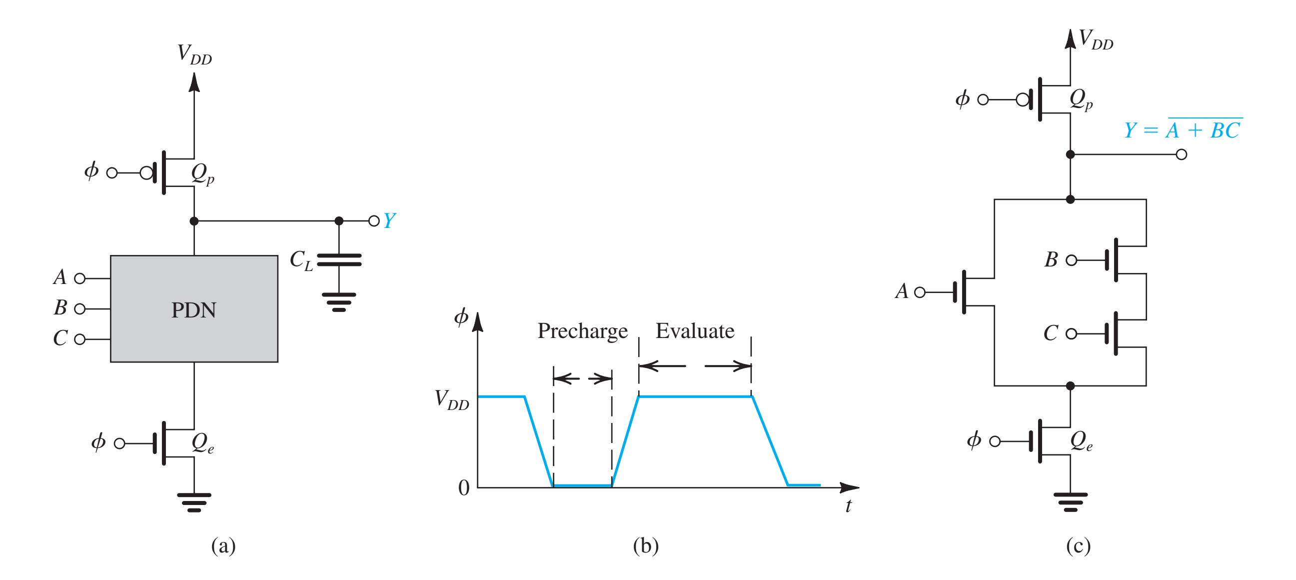


Figure 15.28 **(a)** Basic structure of dynamic MOS logic circuits. **(b)** Waveform of the clock needed to operate the dynamic logic circuit. **(c)** An example circuit.

of static CMOS. For Q_p, we select a W/L ratio large enough to ensure that C_L will be fully charged during the precharge interval, but small enough so that the capacitance C_L will not be increased significantly. This is a ratioless form of MOS logic, where the output levels do not depend on the transistors' W/L ratios (unlike pseudo-NMOS, for instance).

Example 15.4

Consider the four-input, dynamic logic NAND gate shown in Fig. 15.29(a). Assume that the gate is fabricated in a 0.18-μm CMOS technology for which $V_{DD} = 1.8$ V, $V_t = 0.5$ V, and $\mu_n C_{ox} = 4\mu_p C_{ox} = 300\ \mu\text{A/V}^2$. To keep C_L small, NMOS devices with $W/L = 0.27\ \mu\text{m}/0.18\ \mu\text{m}$ are used (including transistor Q_e). The PMOS precharge transistor Q_p has $W/L = 0.54\ \mu\text{m}/0.18\ \mu\text{m}$. The total capacitance C_L is found to be 20 fF.

(a) Consider the precharge operation [Fig. 15.29(b)] with the gate of Q_p at 0 V, and assume that at $t = 0$, C_L is fully discharged. Calculate the rise time of the output voltage, defined as the time for v_Y to rise from 10% to 90% of the final voltage V_{DD}.

(b) For $A = B = C = D = 1$, find the value of t_{PHL}.

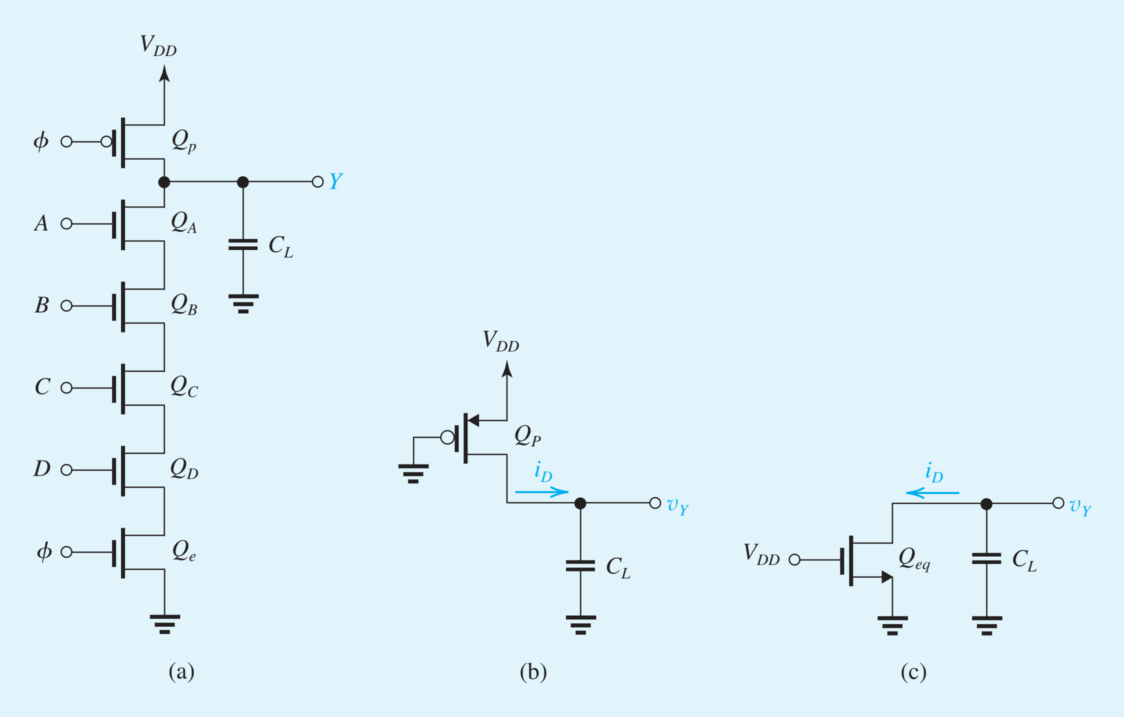


Figure 15.29 Circuits for Example 15.4.

Solution

(a) From Fig. 15.29(b) we see that at $v_Y = 0.1V_{DD} = 0.18$ V, Q_p will be operating in the saturation region and i_D will be

$$i_D(0.1\,V_{DD}) = \frac{1}{2}\mu_p C_{ox}\left(\frac{W}{L}\right)_p (V_{DD} - |V_{tp}|)^2$$

$$= \frac{1}{2} \times 75 \times \frac{0.54}{0.18}(1.8 - 0.5)^2$$

$$= 190.1\ \mu\text{A}$$

At $v_Y = 0.9\,V_{DD} = 1.62$ V, Q_p will be operating in the triode region; thus,

$$i_D(0.9V_{DD}) = \mu_p C_{ox}\left(\frac{W}{L}\right)_p \left[(V_{DD} - |V_{tp}|)\,(V_{DD} - 0.9\,V_{DD}) - \frac{1}{2}(V_{DD} - 0.9V_{DD})^2\right]$$

$$= 75 \times \frac{0.54}{0.18}\left[(1.8 - 0.5)(1.8 - 1.62) - \frac{1}{2}(1.8 - 1.62)^2\right]$$

$$= 49\ \mu\text{A}$$

Thus the average capacitor charging current is

$$I_{\text{av}} = \frac{1}{2}(190.1 + 49) = 119.6\ \mu\text{A}$$

The rise time t_r of v_Y can now be determined from

$$t_r = \frac{C\Delta v_Y}{I_{\text{av}}}$$

$$= \frac{C(0.9V_{DD} - 0.1V_{DD})}{I_{\text{av}}}$$

Thus,

$$t_r = \frac{20 \times 10^{-15} \times 0.8 \times 1.8}{119.6 \times 10^{-6}} = 0.24\ \text{ns}$$

(b) When $A = B = C = D = 1$, all the NMOS transistors will be conducting during the evaluation phase. Replacing the five identical transistors with an equivalent device Q_{eq} with $(W/L)_{\text{eq}} = \frac{1}{5}(W/L) = \frac{1}{5} \times 1.5 = 0.3$, we obtain the equivalent circuit for the capacitor discharge, shown in Fig. 15.29(c). At $v_Y = V_{DD}$, Q_{eq} will be operating in saturation; thus,

$$i_D(V_{DD}) = \frac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)_{\text{eq}} (V_{DD} - V_t)^2$$

$$= \frac{1}{2} \times 300 \times 0.3(1.8 - 0.5)^2$$

$$= 76.1\ \mu\text{A}$$

Example 15.4 *continued*

At $v_Y = V_{DD}/2$, Q_{eq} will be operating in the triode region, thus,

$$i_D(V_{DD}/2) = (\mu_n C_{ox})\left(\frac{W}{L}\right)_{eq}\left[(V_{DD} - V_t)\frac{V_{DD}}{2} - \frac{1}{2}\left(\frac{V_{DD}}{2}\right)^2\right]$$

$$= 300 \times 0.3\left[(1.8 - 0.5)\left(\frac{1.8}{2}\right) - \frac{1}{2}\left(\frac{1.8}{2}\right)^2\right]$$

$$= 68.9\ \mu\text{A}$$

Thus the average capacitor-discharge current is

$$I_{av} = \frac{76.1 + 68.9}{2} = 72.5\ \mu\text{A}$$

and t_{PHL} can be found from

$$t_{PHL} = \frac{C(V_{DD} - V_{DD}/2)}{I_{av}}$$

$$= \frac{20 \times 10^{-15}(1.8 - 0.9)}{72.5 \times 10^{-6}} = 0.25\ \text{ns}$$

EXERCISE

15.14 In an attempt to reduce t_{PHL} of the NAND gate in Example 15.4, the designer doubles the value of W/L of each of the NMOS devices. If C increases to 30 fF, what is the new value of t_{PHL}?
Ans. 0.19 ns

15.5.2 Nonideal Effects

We now briefly consider various sources of nonideal operation of dynamic logic circuits.

Noise Margins Since, during the evaluation phase, the NMOS transistors begin to conduct for $v_I = V_{tn}$,

$$V_{IL} \simeq V_{IH} \simeq V_{tn}$$

GRAND-SCALE GRAPHICS:

IC chips for specialized graphics processing achieve new levels of integration. Among recent announcements from Nividia is a graphics chip (GPU) incorporating 7.1 billion MOS transistors on a 551-mm^2 die in 28-nm CMOS technology from Taiwan Semiconductor Manufacturing Company (TSMC). Besides gaming and graphics, live-streaming video and other applications abound.

and thus the noise margins will be

$$NM_L = V_{IL} - V_{OL} = V_{tn} - 0 = V_{tn}$$
$$NM_H = V_{OH} - V_{IH} = V_{DD} - V_{tn}$$

Thus the noise margins are far from equal, and NM_L is rather low. Although NM_H is high, other nonideal effects reduce its value, as we shall shortly see. At this time, however, observe that the output node is a high-impedance node and thus will be susceptible to noise pickup and other disturbances.

Output Voltage Decay due to Leakage Effects In the absence of a path to ground through the PDN, the output voltage will ideally remain high at V_{DD}. This, however, is based on the assumption that the charge on C_L will remain intact. In practice, there will be leakage current that will cause C_L to slowly discharge and v_Y to decay. The principal source of leakage is the reverse current of the reverse-biased junction between the drain diffusion of transistors connected to the output node and the substrate. Such currents can be in the range of 10^{-12} A to 10^{-15} A, and they increase rapidly with temperature (approximately doubling for every 10°C rise in temperature). Thus the circuit can malfunction if the clock is operating at a very low frequency and the output node is not "refreshed" periodically. This exact same point will be encountered when we study dynamic memory cells in Chapter 16.

Charge Sharing There is another and often more serious way for C_L to lose some of its charge and thus cause v_Y to fall significantly below V_{DD}. To see how this can happen, refer to Fig. 15.30(a), which shows only Q_1 and Q_2, the two top transistors of the PDN, together with the precharge transistor Q_p. Here, C_1 is the capacitance between the common node of Q_1 and Q_2 and ground. At the beginning of the evaluation phase, after Q_p has turned off and with C_L charged to V_{DD} [Fig. 15.30(a)], we assume that C_1 is initially discharged and that the inputs are such that at the gate of Q_1 we have a high signal, whereas at the gate of Q_2 the signal is low. We can easily see that Q_1 will turn on and its drain current, i_{D1}, will flow as indicated. Thus i_{D1} will discharge C_L and charge C_1. Although eventually i_{D1} will reduce to zero, C_L will have lost some of its charge, which will have been transferred to C_1. This phenomenon is known as charge sharing (see Problem 15.49).

We shall not pursue the problem of charge sharing any further here, except to point out a couple of the techniques usually employed to minimize its effect. One approach involves adding a *p*-channel device that continuously conducts a small current to replenish the charge lost by C_L, as shown in Fig. 15.30(b). This arrangement should remind us of pseudo-NMOS. Indeed, adding this transistor will cause the gate to dissipate static power. On the positive side, however, the added transistor will lower the impedance level of the output node and make it less susceptible to noise and will solve the leakage and charge-sharing problems. Another approach to solving the charge-sharing problem is to precharge the internal nodes: that is, to

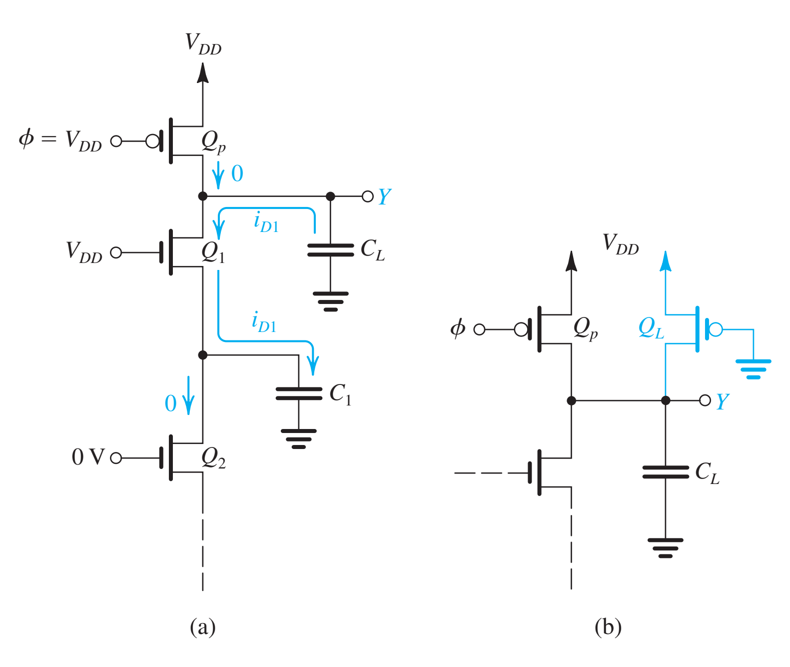


Figure 15.30 **(a)** Charge sharing. **(b)** Adding a permanently turned-on transistor Q_L solves the charge-sharing problem at the expense of static power dissipation.

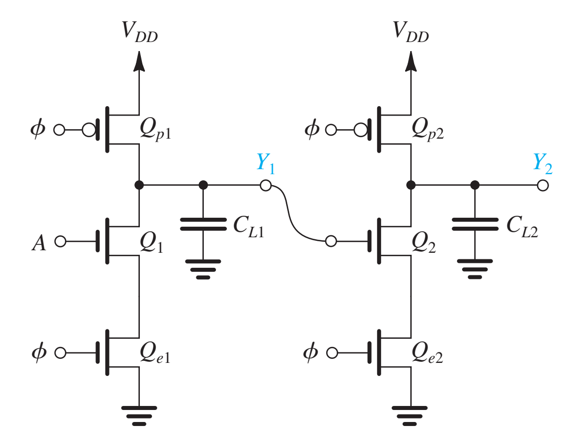


Figure 15.31 Two single-input dynamic logic gates connected in cascade. With the input A high, during the evaluation phase C_{L2} will partially discharge and the output at Y_2 will fall lower than V_{DD}, which can cause logic malfunction.

precharge capacitor C_1. The price paid in this case is increased circuit complexity and node capacitances.

Clock Feedthrough Another problem can arise when the PDN remains off during the evaluation phase. As ϕ rises and turns off Q_p, the output node Y becomes a floating node. However, Y is capacitively coupled to ϕ through C_{gd} of Q_p, and hence the clock signal ϕ can cause a slight rise in output voltage.

Cascading Dynamic Logic Gates A serious problem arises if one attempts to cascade dynamic logic gates. Consider the situation depicted in Fig. 15.31, where two single-input dynamic gates are connected in cascade. During the precharge phase, C_{L1} and C_{L2} will be charged through Q_{p1} and Q_{p2}, respectively. Thus, at the end of the precharge interval, $v_{Y1} = V_{DD}$ and $v_{Y2} = V_{DD}$. Now consider what happens in the evaluation phase for the case of high input A. Obviously, the correct result will be Y_1 low ($v_{Y1} = 0$ V) and Y_2 high ($v_{Y2} = V_{DD}$). What happens, however, is somewhat different. As the evaluation phase begins, Q_1 turns on and C_{L1} begins to discharge. However, simultaneously, Q_2 turns on and C_{L2} also begins to discharge. Only when v_{Y1} drops below V_{tn} will Q_2 turn off. Unfortunately, however, by that time, C_{L2} will have lost a significant amount of its charge, and v_{Y2} will be less than the expected value of V_{DD}. (Here it is important to note that in dynamic logic, once charge has been lost, it cannot be recovered.) This problem is sufficiently serious to make simple cascading an impractical proposition. As usual, however, the ingenuity of circuit designers has come to the rescue, and a number of schemes have been proposed to make cascading possible in dynamic logic circuits. We shall discuss one such scheme after considering Exercise 15.15.

EXERCISE

15.15 To gain further insight into the cascading problem described above, let us determine the decrease in the output voltage v_{Y2} for the circuit in Fig. 15.31. Specifically, consider the circuit as the evaluation phase begins: At $t = 0$, $v_{Y1} = v_{Y2} = V_{DD}$ and $v_{\phi} = v_A = V_{DD}$. Transistors Q_{p1} and Q_{p2} are cut off and can be removed from the equivalent circuit. Furthermore, for the purpose of this approximate analysis, we can replace the series combination of Q_1 and Q_{e1} with a single device having an appropriate W/L, and similarly for the combination of Q_2 and Q_{e2}. The result is the approximate equivalent circuit in Fig. E15.15. We are interested in the operation of this circuit in the interval Δt during which v_{Y1} falls from V_{DD} to V_t, at which time Q_{eq2} turns off and C_{L2} stops discharging. Assume that the process technology has the parameter values specified in Example 15.4 and that for all NMOS transistors in the circuit of Fig. 15.31, $W/L = 4\ \mu\text{m}/2\ \mu\text{m}$ and $C_{L1} = C_{L2} = 40$ fF.

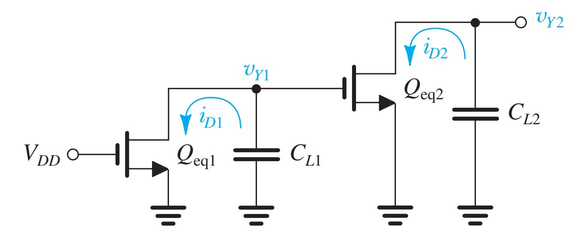


Figure E15.15

(a) Find $(W/L)_{eq1}$ and $(W/L)_{eq2}$.
(b) Find the values of i_{D1} at $v_{Y1} = V_{DD}$ and at $v_{Y1} = V_t$. Hence determine an average value for i_{D1}.
(c) Use the average value of i_{D1} found in (b) to determine an estimate for the interval Δt.

(d) Find the average value of i_{D2} during Δt. To simplify matters, take the average to be the value of i_{D2} obtained when the gate voltage v_{Y1} is midway through its excursion (i.e., $v_{Y1} = 3$ V). (*Hint:* Q_{eq2} will remain in saturation.)

(e) Use the value of Δt found in (c) together with the average value of i_{D2} determined in (d) to find an estimate of the reduction in v_{Y2} during Δt. Hence determine the final value of v_{Y2}.

Ans. (a) 1, 1; (b) 400 μA and 175 μA, for an average value of 288 μA; (c) 0.56 ns; (d) 100 μA; (e) $\Delta v_{Y2} = 1.4$ V, thus v_{Y2} decreases to 3.6 V

15.5.3 Domino CMOS Logic

Domino CMOS logic is a form of dynamic logic that results in cascadable gates. Figure 15.32 shows the structure of the Domino CMOS logic gate. We observe that it is simply the basic dynamic logic gate of Fig. 15.28(a) with a static CMOS inverter connected to its output. Operation of the gate is straightforward. During precharge, X will be raised to V_{DD}, and the gate output Y will be at 0 V. During evaluation, depending on the combination of input variables, either X will remain high and thus the output Y will remain low ($t_{PHL} = 0$) or X will be brought down to 0 V and the output Y will rise to V_{DD} (t_{PLH} finite). Thus, during evaluation, the output either remains low or makes only one low-to-high transition.

To see why Domino CMOS gates can be cascaded, consider the situation in Fig. 15.33(a), where we show two Domino gates connected in cascade. For simplicity, we show single-input gates. At the end of precharge, X_1 will be at V_{DD}, Y_1 will be at 0 V, X_2 will be at V_{DD}, and Y_2 will be at 0 V. As in the preceding case, assume that A is high at the beginning of evaluation. Thus, as ϕ goes up, capacitor C_{L1} will begin discharging, pulling X_1 down. Meanwhile, the low input at the gate of Q_2 keeps Q_2 off, and C_{L2} remains fully charged. When v_{X1} falls below the threshold voltage of inverter I_1, Y_1 will go up, turning Q_2 on, which in turn begins to discharge C_{L2} and pulls X_2 low. Eventually, Y_2 rises to V_{DD}.

From this description, we see that because the output of the Domino gate is low at the beginning of evaluation, no premature capacitor discharge will occur in the subsequent gate in

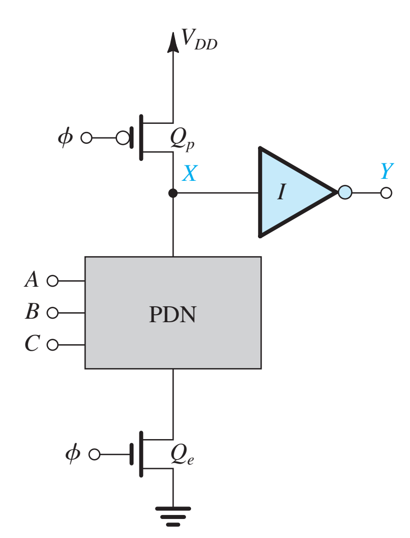


Figure 15.32 The Domino CMOS logic gate. The circuit consists of a dynamic MOS logic gate with a static CMOS inverter connected to the output. During evaluation, Y either will remain low (at 0 V) or will make one 0-to-1 transition (to V_{DD}).

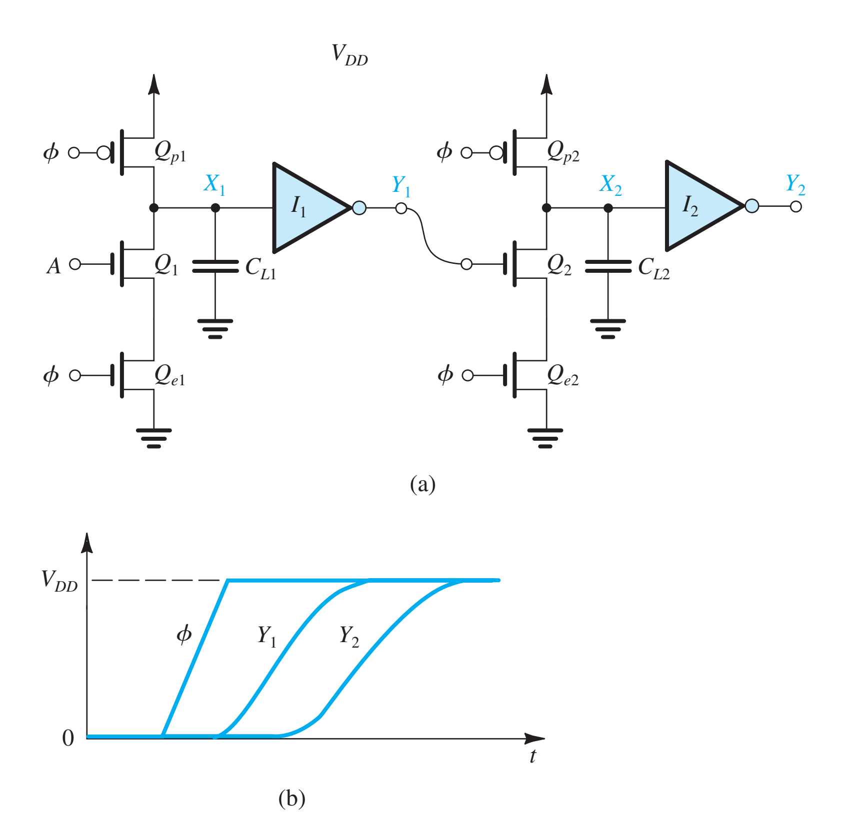


Figure 15.33 **(a)** Two single-input Domino CMOS logic gates connected in cascade. **(b)** Waveforms during the evaluation phase.

the cascade. As indicated in Fig. 15.33(b), output Y_1 will make a 0-to-1 transition t_{PLH} seconds after the rising edge of the clock. Subsequently, output Y_2 makes a 0-to-1 transition after another t_{PLH} interval. The propagation of the rising edge through a cascade of gates resembles contiguously placed dominoes falling over, each toppling the next, which is the origin of the name Domino CMOS logic. Domino CMOS logic finds application in the design of address decoders in memory chips, for example.

15.5.4 Concluding Remarks

Dynamic logic presents many challenges to the circuit designer. Although it can provide high-speed operation, as well as considerable reduction in the chip-area requirement and zero (or little) static power dissipation, the circuits are prone to many nonideal effects, some of which have been discussed here. It should also be remembered that dynamic power dissipation is an important issue in dynamic logic. Another factor that should be considered is the "dead time" during precharge when the output of the circuit is not yet available.

15.6 Bipolar and BiCMOS Logic Circuits

A

As mentioned earlier, the use of the BJT in designing digital logic circuits has diminished considerably. At the present time, only one type of bipolar logic circuit is still in use, albeit

for specialized applications. This is emitter-coupled logic (ECL), which we very briefly study here. As well, we shall take a brief look at a digital circuit technology that combines bipolar and CMOS and is aptly named BiCMOS.

15.6.1 Emitter-Coupled Logic (ECL)

Emitter-coupled logic (ECL) is the fastest logic-circuit family available for conventional logic-system design.[6] High speed is achieved by operating all bipolar transistors out of saturation, thus avoiding storage-time delays, and by keeping the logic signal swings relatively small (about 0.8 V or less), thus reducing the time required to charge and discharge the various load and parasitic capacitances. Saturation in ECL is avoided by using the BJT differential pair as a current switch. The BJT differential pair was studied in Chapter 9, and we urge the reader to review the introduction given in Section 9.2 before proceeding with the study of ECL.

The Basic Principle Emitter-coupled logic is based on the use of the current-steering switch introduced in Section 14.2 (Fig. 14.19). Such a switch can be most conveniently realized using the differential pair shown in Fig. 15.34. The pair is biased with a constant-current source I, and one side is connected to a reference voltage V_R. As shown in Section 9.2, the current I can be steered to either Q_1 or Q_2 under the control of the input signal v_I. Specifically, when v_I is greater than V_R by about $4V_T$ ($\simeq$100 mV), nearly all the current I is conducted by Q_1, and thus for $\alpha_1 \simeq 1$, $v_{O1} = V_{CC} - IR_C$. Simultaneously, the current through Q_2 will be nearly zero, and thus $v_{O2} = V_{CC}$. Conversely, when v_I is lower than V_R by about $4V_T$, most of the current

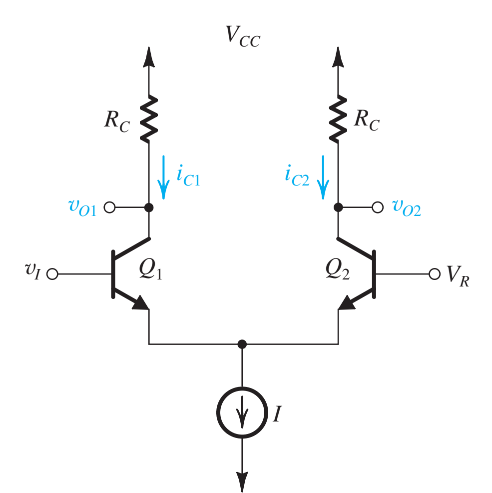


Figure 15.34 The basic element of ECL is the differential pair. Here, V_R is a reference voltage.

[6]Higher speeds of operation can be obtained with gallium arsenide (GaAs) circuits; these, however, are not available as off-the-shelf components for conventional digital system design. GaAs digital circuits are not covered in this book; however, a substantial amount of material on this subject can be found on the book's website.

I will flow through Q_2 and the current through Q_1 will be nearly zero. Thus $v_{O1} = V_{CC}$ and $v_{O2} = V_{CC} - IR_C$.

The preceding description suggests that as a logic element, the differential pair realizes an inversion function at v_{O1} and simultaneously provides the complementary output signal at v_{O2}. The output logic levels are $V_{OH} = V_{CC}$ and $V_{OL} = V_{CC} - IR_C$, and thus the output logic swing is IR_C. A number of additional remarks can be made concerning this circuit:

1. The differential nature of the circuit makes it less susceptible to picked-up noise. In particular, an interfering signal will tend to affect both sides of the differential pair similarly and thus will not result in current switching. This is the common-mode rejection property of the differential pair (see Section 9.2).
2. The current drawn from the power supply remains constant during switching. Thus, unlike CMOS (and TTL), no supply current spikes occur in ECL, eliminating an important source of noise in digital circuits. This is a definite advantage, especially since ECL is usually designed to operate with small signal swings and has correspondingly low noise margins.
3. The output signal levels are both referenced to V_{CC} and thus can be made particularly stable by operating the circuit with $V_{CC} = 0$: in other words, by utilizing a negative power supply and connecting the V_{CC} line to ground. In this case, $V_{OH} = 0$ and $V_{OL} = -IR_C$.
4. Some means must be provided to make the output signal levels compatible with those at the input so that one gate can drive another. Practical ECL gate circuits incorporate a level-shifting arrangement that serves to center the output signal levels on the value of V_R.
5. The availability of complementary outputs considerably simplifies logic design with ECL.

EXERCISE

15.16 For the circuit in Fig. 15.34, let $V_{CC} = 0$, $I = 4$ mA, $R_C = 220\,\Omega$, $V_R = -1.32$ V, and assume $\alpha \simeq 1$. Determine V_{OH} and V_{OL}. By how much should the output levels be shifted so that the values of V_{OH} and V_{OL} become centered on V_R? What will the shifted values of V_{OH} and V_{OL} be?
Ans. 0; –0.88 V; –0.88 V; –0.88 V, –1.76 V

The Basic Gate Circuit The basic gate circuit of a particular ECL logic-circuit family, known as ECL 10K, is shown in Fig. 15.35. The circuit consists of three parts. The network composed of Q_1, D_1, D_2, R_1, R_2, and R_3 generates a reference voltage V_R whose value at room temperature is –1.32 V. It can be shown that the value of this reference voltage changes with temperature in a predetermined manner, keeping the noise margins almost constant. Also, the reference voltage V_R is made relatively insensitive to variations in the power-supply voltage V_{EE}.

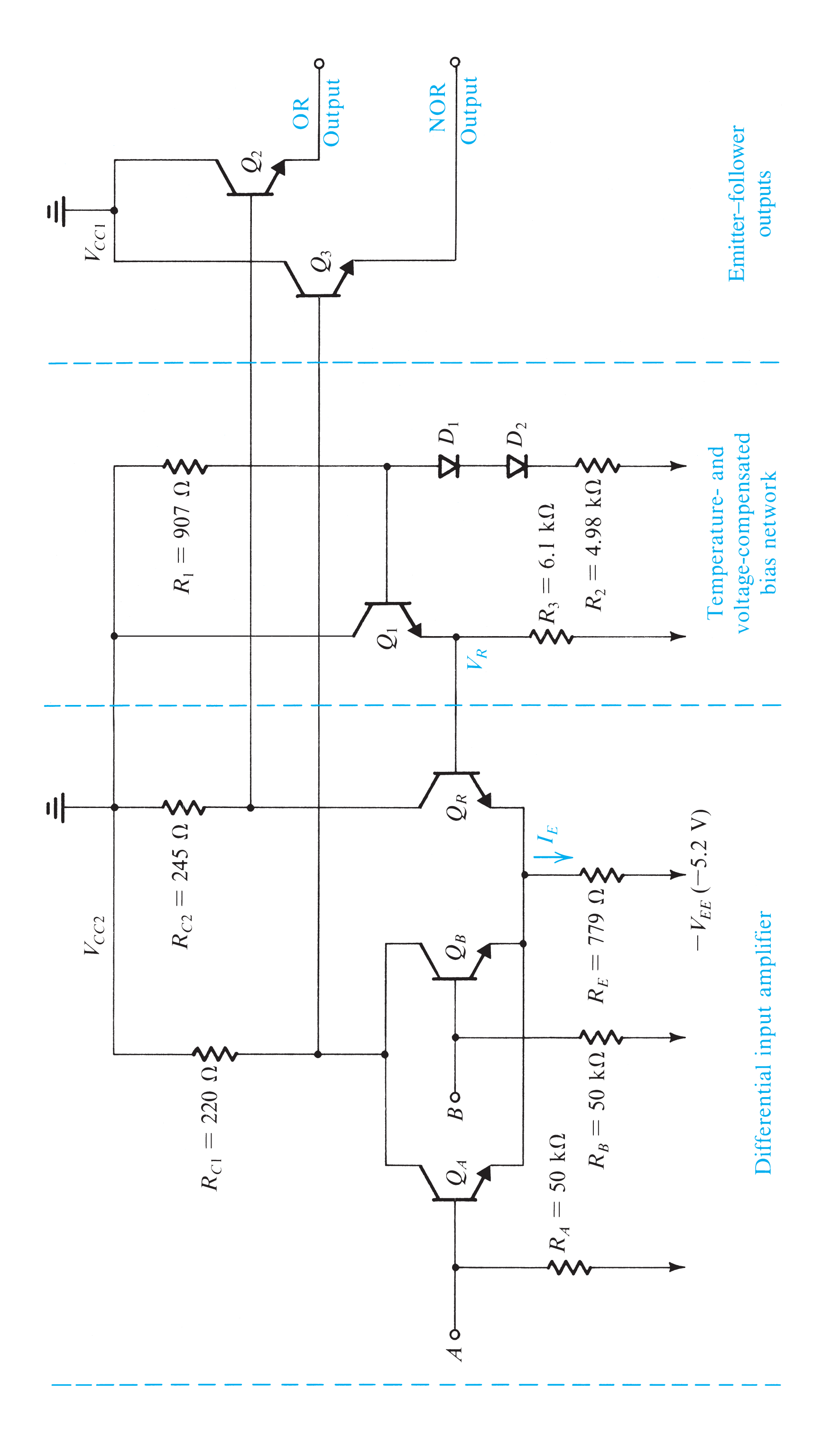


Figure 15.35 Basic circuit of the ECL 10K logic-gate family.

EXERCISE

15.17 Figure E15.17 shows the circuit that generates the reference voltage V_R. Assuming that the voltage drop across each of D_1, D_2, and the base–emitter junction of Q_1 is 0.75 V, calculate the value of V_R. Neglect the base current of Q_1.

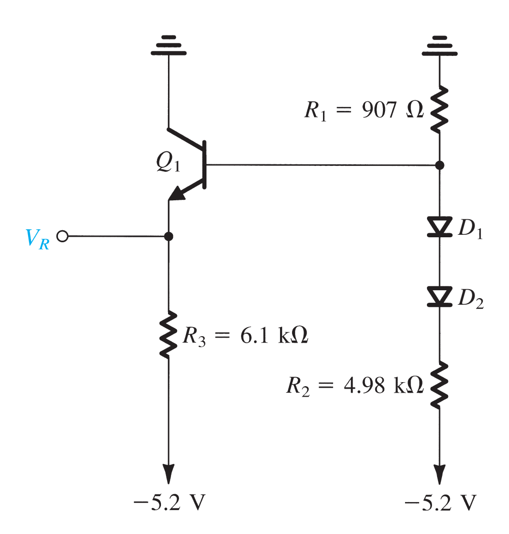


Figure E15.17

Ans. –1.32 V

The second part, and the heart of the gate, is the differential amplifier formed by Q_R and either Q_A or Q_B. This differential amplifier is biased not by a constant-current source, as was done in the circuit of Fig. 15.34, but with a resistance R_E connected to the negative supply $-V_{EE}$. Nevertheless, it can be shown that the current in R_E remains approximately constant over the normal range of operation of the gate. One side of the differential amplifier consists of the reference transistor Q_R, whose base is connected to the reference voltage V_R. The other side consists of a number of transistors (two in the case shown), connected in parallel, with separated bases, each connected to a gate input. If the voltages applied to A and B are at the logic-0 level, which is about 0.4 V below V_R, both Q_A and Q_B will be off and the current I_E in R_E will flow through the reference transistor Q_R. The resulting voltage drop across R_{C2} will cause the collector voltage of Q_R to be low.

On the other hand, when the voltage applied to A or B is at the logic-1 level, which is about 0.4 V above V_R, transistor Q_A or Q_B, or both, will be on and Q_R will be off. Thus the current I_E will flow through Q_A or Q_B, or both, and an almost equal current will flow through R_{C1}. The resulting voltage drop across R_{C1} will cause the collector voltage to drop. Meanwhile, since Q_R is off, its collector voltage rises. We thus see that the voltage at the collector of Q_R will be high if A or B, or both, is high, and thus at the collector of Q_R, the OR logic function, $A + B$, is realized. On the other hand, the common collector of Q_A and Q_B will be high only when A and B are simultaneously low. Thus at the common collector of Q_A and Q_B, the logic function

$\overline{A}\,\overline{B} = \overline{A+B}$ is realized. We therefore conclude that the two-input gate of Fig. 15.35 realizes the OR function and its complement, the NOR function. The availability of complementary outputs is an important advantage of ECL; it simplifies logic design and avoids the use of additional inverters with associated time delay.

It should be noted that the resistance connecting each of the gate input terminals to the negative supply enables the user to leave an unused input terminal open: An open input terminal will be *pulled down* to the negative supply voltage, and its associated transistor will be off.

EXERCISE

15.18 With input terminals A and B in Fig. 15.35 left open, find the current I_E through R_E. Also find the voltages at the collector of Q_R and at the common collector of the input transistors Q_A and Q_B. Use $V_R = -1.32$ V, V_{BE} of $Q_R \simeq 0.75$ V, and assume that β of Q_R is very high.
Ans. 4 mA; –1 V; 0 V

The third part of the ECL gate circuit is composed of the two emitter followers, Q_2 and Q_3. The emitter followers do not have on-chip loads, since in many applications of high-speed logic circuits the gate output drives a transmission line terminated at the other end, as indicated in Fig. 15.36.

The emitter followers have two purposes: First, they shift the level of the output signals by one V_{BE} drop. Thus, using the results of Exercise 15.18, we see that the output levels become approximately –1.75 V and –0.75 V. These shifted levels are centered approximately around the reference voltage ($V_R = -1.32$ V), which means that one gate can drive another. This compatibility of logic levels at input and output is an essential requirement in the design of gate circuits.

The second function of the output emitter followers is to provide the gate with low output resistances and with the large output currents required for charging load capacitances. Since these large transient currents can cause spikes on the power-supply line, the collectors of the emitter followers are connected to a power-supply terminal V_{CC1} separate from that of the

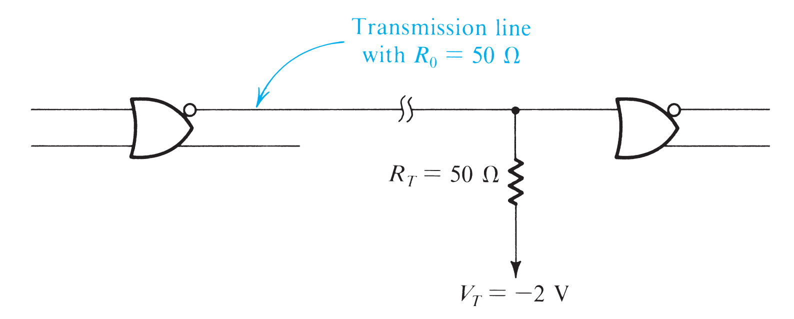


Figure 15.36 The proper way to connect high-speed logic gates such as ECL. Properly terminating the transmission line connecting the two gates eliminates the "ringing" that would otherwise corrupt the logic signals.

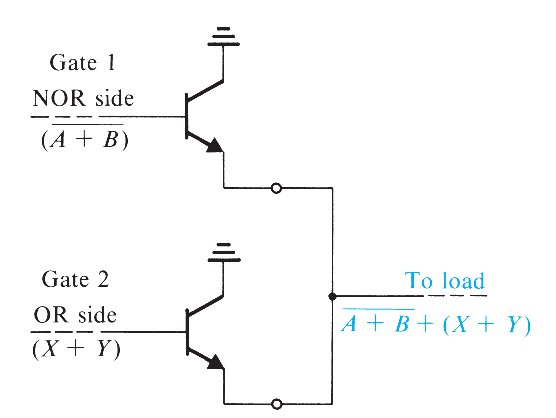


Figure 15.37 The wired-OR capability of ECL.

differential amplifier and the reference-voltage circuit, V_{CC2}. Here we note that the supply current of the differential amplifier and the reference circuit remains almost constant. The use of separate power-supply terminals prevents the coupling of power-supply spikes from the output circuit to the gate circuit and thus lessens the likelihood of false gate switching. Both V_{CC1} and V_{CC2} are of course connected to the same system ground, external to the chip.

The Wired-OR Capability The emitter–follower output stage of the ECL family allows an additional level of logic to be performed at very low cost by simply wiring the outputs of several gates in parallel. This is illustrated in Fig. 15.37, where the outputs of two gates are wired together. Note that the base–emitter diodes of the output followers realize an OR function: This **wired-OR** connection can be used to provide gates with high fan-in as well as to increase the flexibility of ECL in logic design.

Final Comments We have chosen to study ECL by focusing on a commercially available circuit family. A great deal of design optimization has been applied to create a very-high-performance family of SSI and MSI logic circuits. ECL and some of its variants are also used in VLSI circuit design. Applications include very-high-speed processors such as those used in supercomputers, as well as high-speed and high-frequency communication systems. When employed in VLSI design, current-source biasing is almost always utilized. Further, a variety of circuit configurations are employed (see Rabaey, 1996). The major disadvantage of ECL is its large static power dissipation, which limits its use to specialized applications where high speed of operation is of paramount interest.

15.6.2 BiCMOS Digital Circuits

In this section, we provide an introduction to a VLSI circuit technology that combines *bi*polar and *CMOS* circuits on one IC chip. The aim is to combine the low-power, high-input impedance and wide noise margins of CMOS with the high current-driving capability of bipolar transistors. Specifically, CMOS, although a nearly ideal logic-circuit technology in many respects, has a limited current-driving capability. This is not a serious problem when the CMOS gate has to drive a few other CMOS gates. It becomes a serious issue, however, when relatively large capacitive loads (e.g., greater than 0.5 pF or so) are present. In such cases, one has to either resort to the use of elaborate CMOS buffer circuits (such as those discussed in Section 14.4.4) or face the usually unacceptable consequence of long propagation delays. On the other hand, we know that by virtue of its much larger transconductance, the BJT is capable of large output currents. We have seen a practical illustration of that in the emitter–follower

output stage of ECL. Indeed, the high current-driving capability contributes to making ECL two to five times faster than CMOS (under equivalent conditions)—of course, at the expense of high power dissipation. In summary, then, BiCMOS seeks to combine the best of the CMOS and bipolar technologies to obtain a class of circuits that is particularly useful when output currents that are higher than is possible with CMOS are needed. The price paid is a processing technology that is more complex, and hence more expensive, than CMOS.

The BiCMOS Inverter A variety of BiCMOS inverter circuits have been proposed and are in use. All of these are based on the use of *npn* transistors to increase the output current available from a CMOS inverter. This can be most simply achieved by cascading each of the Q_N and Q_P devices of the CMOS inverter with an *npn* transistor, as shown in Fig. 15.38(a). Observe that this circuit can be thought of as utilizing the pair of complementary composite MOS-BJT devices shown in Fig. 15.38(b). These composite devices[7] retain the high input impedance of the MOS transistor while in effect multiplying its rather low g_m by the β of the BJT. It is also useful to observe that the output stage formed by Q_1 and Q_2 has what is known as the **totem-pole configuration** utilized by TTL.[8]

The circuit of Fig. 15.38(a) operates as follows: When v_I is low, both Q_N and Q_2 are off while Q_P conducts and supplies Q_1 with base current, thus turning it on. Transistor Q_1 then provides a large output current to charge the load capacitance. The result is a very fast charging of the load capacitance and correspondingly a short low-to-high propagation delay, t_{PLH}. Transistor Q_1 turns off when v_O reaches a value of about $V_{DD} - V_{BE1}$, and thus the output high level is lower than V_{DD}, a disadvantage. When v_I goes high, Q_P and Q_1 turn off, and Q_N turns on, providing its drain current into the base of Q_2. Transistor Q_2 then turns on and provides a large output current that quickly discharges the load capacitance. Here again the result is a short high-to-low propagation delay, t_{PHL}. On the negative side, Q_2 turns off when v_O reaches a value of about V_{BE2}, and thus the output low level is greater than zero, a disadvantage.

Thus, while the circuit of Fig. 15.38(a) features large output currents and short propagation delays, it has the disadvantage of reduced logic swing and, correspondingly, reduced noise margins. There is also another and perhaps more serious disadvantage, namely, the relatively long turn-off delays of Q_1 and Q_2 arising from the absence of circuit paths along which the base charge can be removed. This problem can be solved by adding a resistor between the base of each of Q_1 and Q_2 and ground, as shown in Fig. 15.38(c). Now when either Q_1 or Q_2 is turned off, its stored base charge is removed to ground through R_1 or R_2, respectively. Resistor R_2 provides an additional benefit: With v_I high, and after Q_2 cuts off, v_O continues to fall below V_{BE2}, and the output node is pulled to ground through the series path of Q_N and R_2. Thus R_2 functions as a pull-down resistor. The Q_N–R_2 path, however, is a high-impedance one with the result that pulling v_O to ground is a rather slow process. Incorporating the resistor R_1, however, is disadvantageous from a static power-dissipation standpoint: When v_I is low, a dc path exists between V_{DD} and ground through the conducting Q_P and R_1. Finally, it should be noted that R_1 and R_2 take some of the drain currents of Q_P and Q_N away from the bases of

[7]It is interesting to note that these composite devices were proposed as early as 1969 (see Lin et al., 1969).

[8]Refer to the book's website for a description of the basic TTL logic-gate circuit and its totem-pole output stage.

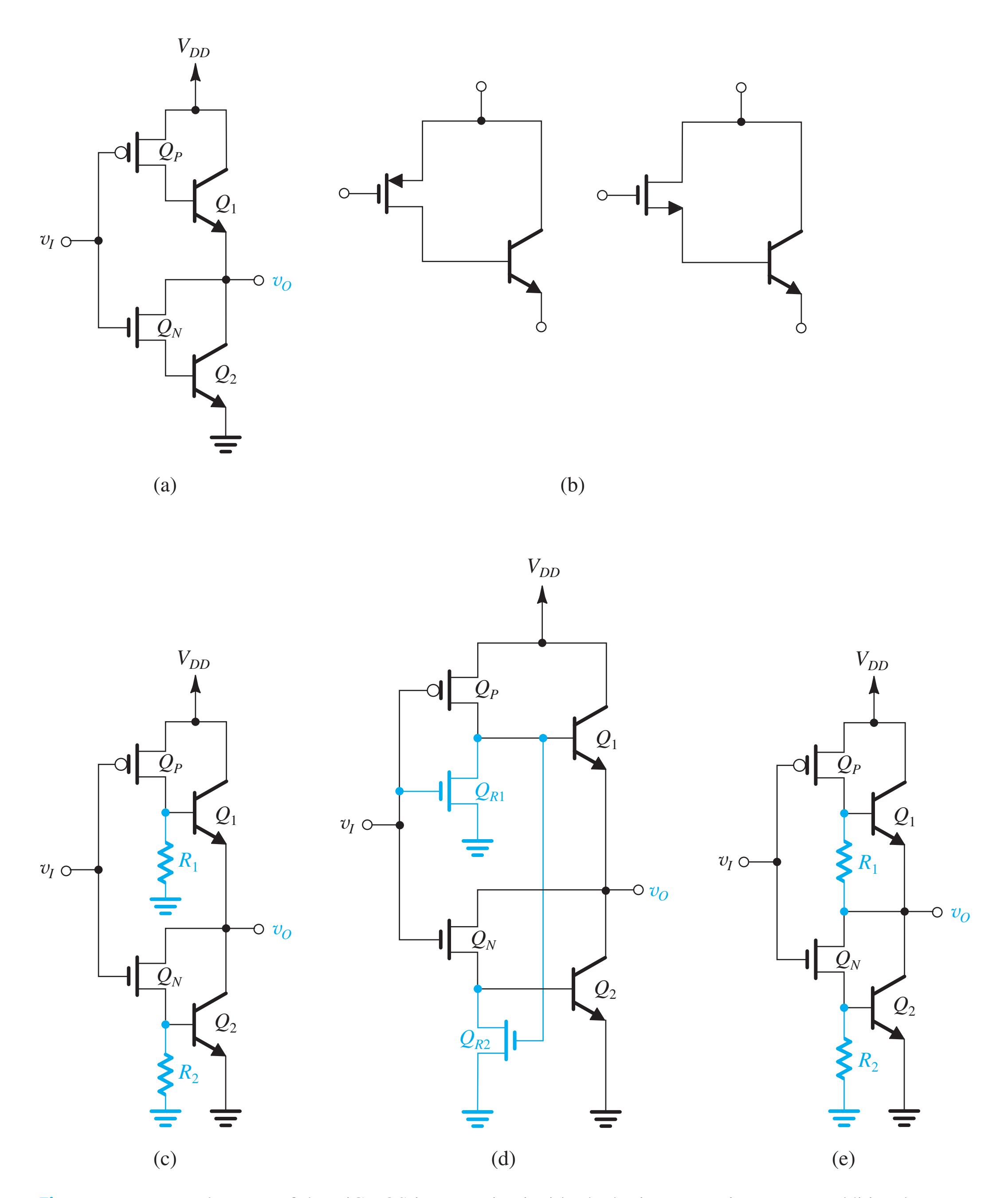


Figure 15.38 Development of the BiCMOS inverter circuit. **(a)** The basic concept is to use an additional bipolar transistor to increase the output current drive of each of Q_N and Q_P of the CMOS inverter. **(b)** The circuit in **(a)** can be thought of as utilizing these composite devices. **(c)** To reduce the turn-off times of Q_1 and Q_2, "bleeder resistors" R_1 and R_2 are added. **(d)** Implementation of the circuit in **(c)** using NMOS transistors to realize the resistors. **(e)** An improved version of the circuit in **(c)** obtained by connecting the lower end of R_1 to the output node.

Q_1 and Q_2 and thus slightly reduce the gate output current available to charge and discharge the load capacitance.

Figure 15.38(d) shows the way in which R_1 and R_2 are usually implemented. As indicated, NMOS devices Q_{R1} and Q_{R2} are used to realize R_1 and R_2. As an added innovation, these two transistors are made to conduct only when needed. Thus, Q_{R1} will conduct only when v_I rises, at which time its drain current constitutes a reverse base current for Q_1, speeding up its turn-off. Similarly, Q_{R2} will conduct only when v_I falls and Q_P conducts, pulling the gate of Q_{R2} high. The drain current of Q_{R2} then constitutes a reverse base current for Q_2, speeding up its turn-off.

As a final circuit for the BiCMOS inverter, we show the so-called R-circuit in Fig. 15.38(e). This circuit differs from that in Fig. 15.38(c) in only one respect: Rather than returning R_1 to ground, we have connected R_1 to the output node of the inverter. This simple change has two benefits. First, the problem of static power dissipation is now solved. Second, R_1 now functions as a pull-up resistor, pulling the output node voltage up to V_{DD} (through the conducting Q_P) after Q_1 has turned off. Thus, the R circuit in Fig. 15.38(e) does in fact have output levels very close to V_{DD} and ground.

Finally, note that despite the initial promise of BiCMOS, the added processing complexity has somewhat hindered the penetration of BiCMOS into digital IC design. Its use is currently limited to specialized applications, including radio-frequency (RF) circuits that employ an advanced bipolar technology known as silicon–germanium (SiGe).

EXERCISE

D15.19 The threshold voltage of the BiCMOS inverter of Fig. 15.38(e) is the value of v_I at which both Q_N and Q_P are conducting equal currents and operating in the saturation region. At this value of v_I, Q_2 will be on, causing the voltage at the source of Q_N to be approximately 0.7 V. It is required to design the circuit so that the threshold voltage is equal to $V_{DD}/2$. For $V_{DD} = 5$ V, $|V_t| = 0.6$ V, and assuming equal channel lengths for Q_N and Q_P and that $\mu_n \simeq 2.5\,\mu_p$, find the required ratio of widths, W_p/W_n.
Ans. 1

Summary

- The third significant metric in digital IC design, along with speed of operation and power dissipation, is the size of the silicon area required for an inverter.
- Refer to Table 15.1 for the implications of scaling the dimensions of the MOSFET and V_{DD} and V_t by a factor $1/S$.
- In devices with short channels ($L < 0.25$ µm) velocity saturation occurs. Its effect is that i_D saturates early, and its value is lower than would be the case in long-channel devices (see Figs. 15.3, 15.4, and 15.5, and Eq. 15.11).
- Subthreshold conduction is increasingly becoming an important issue in CMOS circuits, leading to significant static power consumption.
- Predominantly because of its low power dissipation and because of its scalability, CMOS is by far the principal technology for digital IC design. This dominance is expected to continue for many years to come.
- Standard CMOS logic utilizes two transistors, an NMOS and a PMOS, for each input variable. Thus the circuit complexity, silicon area, and parasitic capacitance all increase with fan-in.

- To reduce the device count, two other forms of static CMOS, namely, pseudo-NMOS and pass-transistor logic (PTL), are employed in special applications as supplements to standard CMOS.
- Pseudo-NMOS utilizes the same PDN as in standard CMOS logic but replaces the PUN with a single PMOS transistor whose gate is grounded and thus is permanently on. Unlike standard CMOS, pseudo-NMOS is a ratioed form of logic in which V_{OL} is determined by the ratio r of k_n to k_p. Normally, r is selected in the range of 4 to 10 and its value determines the noise margins.
- Pseudo-NMOS has the disadvantage of dissipating static power when the output of the logic gate is low. Static power can be eliminated by turning the PMOS load on for only a brief interval, known as the precharge interval, to charge the capacitance at the output node to V_{DD}. Then the inputs are applied, and depending on the input combination, the output node either remains high or is discharged through the PDN. This is the essence of dynamic logic.
- Pass-transistor logic utilizes either single NMOS transistors or CMOS transmission gates to implement a network of switches that are controlled by the input logic variables. Switches implemented by single NMOS transistors, though simple, result in the reduction of V_{OH} from V_{DD} to $V_{DD} - V_t$.
- The CMOS transmission gate, composed of the parallel connection of an NMOS and a PMOS transistor, is a very effective switch in both analog and digital applications. It passes the entire input signal swing, 0 to V_{DD}. As well, it has an almost constant on-resistance over the full output range.
- A particular form of dynamic logic circuits, known as Domino logic, allows the cascading of dynamic logic gates.
- Emitter-coupled logic (ECL) is the fastest commercially available logic-circuit family. It achieves its high speed of operation by avoiding transistor saturation and by utilizing small logic-signal swings. Its high speed of operation is achieved at the expense of large power dissipation, which limits its application to highly specialized applications.
- The ECL gate provides two complementary outputs, realizing the OR and NOR functions. The outputs of ECL gates can be wired together to realize the OR function of the individual output variables.
- BiCMOS combines the low power and wide noise margins of CMOS with the high current-driving capability (and thus the short gate delays) of BJTs. However, the added processing complexity (over that required for CMOS) has limited its use to specialized applications.

PROBLEMS

Section 15.1: Implications of Technology Scaling: Issues in Deep-Submicron Design

15.1 A chip with a certain area designed using the 5-μm process of the late 1970s contains 20,000 transistors. What does Moore's law predict the number of transistors to be on a chip of equal area fabricated using the 32-nm process of 2013?

15.2 Consider the scaling from a 0.13-μm process to a 65-nm process.

(a) Assuming V_{DD} and V_t are scaled by the same factor as the device dimensions ($S = 2$), find the factor by which t_P, the maximum operating speed, P_{dyn}, power density, and *PDP* decrease (or increase)?

(b) Repeat (a) for the situation in which V_{DD} and V_t remain unchanged.

15.3 For a 65-nm technology, $V_{DS\text{sat}}$ for minimum-length NMOS devices is measured to be 0.25 V and that for minimum-length PMOS devices 0.45 V. What do you estimate the effective values of μ_n and μ_p to be? Also find the values of E_{cr} for both device polarities.

15.4 Consider NMOS and PMOS transistors with minimum channel length fabricated in a 0.13-μm CMOS process. If the effective values of μ_n and μ_p are 350 cm^2/V · s and

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

150 cm^2/V · s, respectively, find the expected values of V_{DSsat} for both device polarities.

15.5 (a) Show that for a short-channel NMOS transistor, the ratio of the current I_{Dsat} obtained at $v_{GS} = V_{DD}$ to the current obtained if velocity saturation were absent is given by

$$\frac{I_{Dsat}}{I_D} = \frac{2V_{DSsat}\left(V_{DD} - V_t - \frac{1}{2}V_{DSsat}\right)}{(V_{DD} - V_t)^2}$$

(b) Find the ratio in (a) for a transistor fabricated in a 65-nm process with $L = 65$-nm, $V_t = 0.35$ V, $V_{DSsat} = 0.25$ V, and $V_{DD} = 1.0$ V.

15.6 (a) Consider a CMOS inverter fabricated in a deep-submicron technology utilizing transistors with the minimum allowed channel length and having an equivalent load capacitance C. Let v_I rise instantaneously to V_{DD} and assume that Q_P turns off and Q_N turns on immediately. Ignoring channel-length modulation, that is, $\lambda = 0$, and assuming Q_N operates in the velocity-saturation region, show that

$$t_{PHL} = \frac{CV_{DD}}{2I_{Dsat}}$$

(b) Using the equivalent resistance of Q_N show that

$$t_{PHL} = 0.69C\frac{12.5 \times 10^3}{(W/L)_n}$$

(c) If the formulas in (a) and (b) are to yield the same result, find V_{DSsat} for the NMOS transistor for a 0.13-µm technology characterized by $V_{DD} = 1.2$ V, $V_t = 0.4$ V, and $\mu_n C_{ox} = 325\ \mu$A/V^2.

D 15.7 (a) For a CMOS inverter fabricated in a deep-submicron technology with $L_n = L_p =$ the minimum allowed channel length, it is required to select W_p/W_n so that $t_{PHL} = t_{PLH}$. This can be achieved by making I_{Dsat} of Q_N equal to I_{Dsat} of Q_P at $|v_{GS}| = V_{DD}$. Show that W_p/W_n is given by

$$\frac{W_p}{W_n} = \frac{\mu_n}{\mu_p}\frac{V_{DSsatn}}{|V_{DSsatp}|}\frac{V_{DD} - V_{tn} - \frac{1}{2}V_{DSsatn}}{V_{DD} - |V_{tp}| - \frac{1}{2}|V_{DSsatp}|}$$

(b) Find the required W_p/W_n for a 65-nm technology for which $\mu_n/\mu_p = 4$, $V_{DD} = 1.0$ V, $V_{tn} = -V_{tp} = 0.35$ V, $V_{DSsatn} = 0.25$ V, and $|V_{DSsatp}| = 0.45$ V.

D 15.8 The current I_S in the subthreshold conduction Eq. (15.13) is proportional to e^{-V_t/nV_T}. If the threshold voltage of an NMOS transistor is reduced by 0.1 V, by what factor will the static power dissipation increase? Assume $n = 2$. Repeat for a reduction in V_t by 0.2 V. What do you conclude about the selection of a value of V_t in process design?

15.9 Measurements on a MOSFET operating in the subthreshold conduction region indicate that the current changes by a factor of 10 for every 80-mV change in v_{GS} and that $i_D = 20$ nA at $v_{GS} = 0.16$ V.

(a) Find the value of i_D at $v_{GS} = 0$.
(b) For a chip having 1 billion transistors, find the current drawn from the 1-V supply V_{DD} as a result of subthreshold conduction. Hence, estimate the resulting static power dissipation.

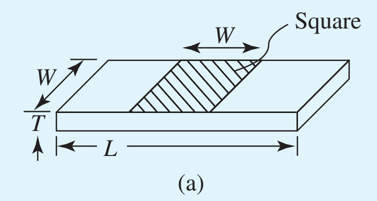
15.10 An NMOS transistor with $k_n = 0.4$ mA/V^2 and a nominal V_{tn} of 0.4 V is to operate in saturation at $I_D = 0.2$ mA.

(a) If V_{tn} can vary by as much as ±10%, what is the expected range of I_D obtained?
(b) If the transistor is used to discharge a 100-fF load capacitance, what is the expected variation in delay time, assuming that the output voltage is to change by 0.1 V?

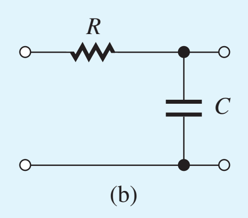
15.11 An interconnect wire with a length L, a width W, and a thickness T has a resistance R given by

$$R = \rho\frac{L}{A} = \frac{\rho L}{TW}$$

where ρ is the resistivity of the material of which the wire is made. The quantity ρ/T is called the **sheet resistance** and has the dimension of ohms, although it is usually expressed as ohms/square or Ω/□ (refer to Fig. P15.11a).



(a)



(b)

Figure P15.11

(a) Find the resistance of an aluminum wire that is 5 mm long and 0.5 μm wide, if the sheet resistance is specified to be 27 mΩ/□.
(b) If the wire capacitance to ground is 0.1 fF/μm length, what is the total wire capacitance?
(c) If we can model the wire very approximately as an RC circuit as shown in Fig. P15.11(b), find the delay time introduced by the wire. (*Hint*: $t_{delay} = 0.69RC$.) (P.S. Only a small fraction of the interconnect on an IC would be this long!)

Section 15.3: Pseudo-NMOS Logic Circuits

15.12 The purpose of this problem is to compare the value of t_{PLH} obtained with a resistive load [see Fig. P15.12(a)] to that obtained with a current-source load [see Fig. P15.12(b)]. For a fair comparison, let the current source $I = V_{DD}/R_D$, which is the initial current available to charge the capacitor in the case of a resistive load. Find t_{PLH} for each case, and hence the percentage reduction obtained when a current-source load is used.

15.13 For a pseudo-NMOS inverter fabricated in a 0.13-μm process and having $k_n = 5k_p = 500\ \mu A/V^2$, $V_{tn} = -V_{tp} = 0.4$ V, and $V_{DD} = 1.3$ V, find V_{OH}, V_{OL}, and I_{sat}. Hence, find the static power dissipation in the low-output state.

15.14 For the pseudo-NMOS inverter specified in Problem 15.13, find V_{OL}, V_{IL}, V_M, V_{IH}, V_{OH}, NM_L, and NM_H.

15.15 Find t_{PLH}, t_{PHL}, and t_P for the pseudo-NMOS inverter specified in Problem 15.13 when loaded with $C = 10$ fF.

D *15.16 Design a pseudo-NMOS inverter that has equal capacitive charging and discharging currents at $v_O = V_{DD}/4$ for use in a system with $V_{DD} = 2.5$ V, $|V_t| = 0.5$ V, $k_n' = 115\,\mu A/V^2$, $k_p' = 30\,\mu A/V^2$, and $(W/L)_n = 1.5$. What are the values of $(W/L)_p$, V_{IL}, V_{IH}, V_M, V_{OH}, V_{OL}, NM_H, and NM_L?

***15.17** Use Eq. (15.26) to find the value of r for which NM_L is maximized. What is the corresponding value of NM_L for the case $V_{DD} = 1.3$ V and $V_t = 0.4$ V? Show that NM_L does not change very much with r by evaluating NM_L for $r = 2$, 5, and 10.

15.18 For what value of r does NM_H of a pseudo-NMOS inverter become zero? Prepare a table of NM_H and NM_L versus r, for $r = 2$ to 10. Let $V_{DD} = 1.3$ V and $V_t = 0.4$ V. Use your table and iteration to determine the value of r that results in $NM_L = NM_H$. What is the resulting margin?

D 15.19 Design a pseudo-NMOS inverter that has $V_{OL} = 0.1$ V. Let $V_{DD} = 1.8$ V, $|V_t| = 0.5$ V, $k_n' = 4k_p' = 400\,\mu A/V^2$, and $(W/L)_p = 1$. What is the value of $(W/L)_n$? Calculate the values of NM_L and the static power dissipation.

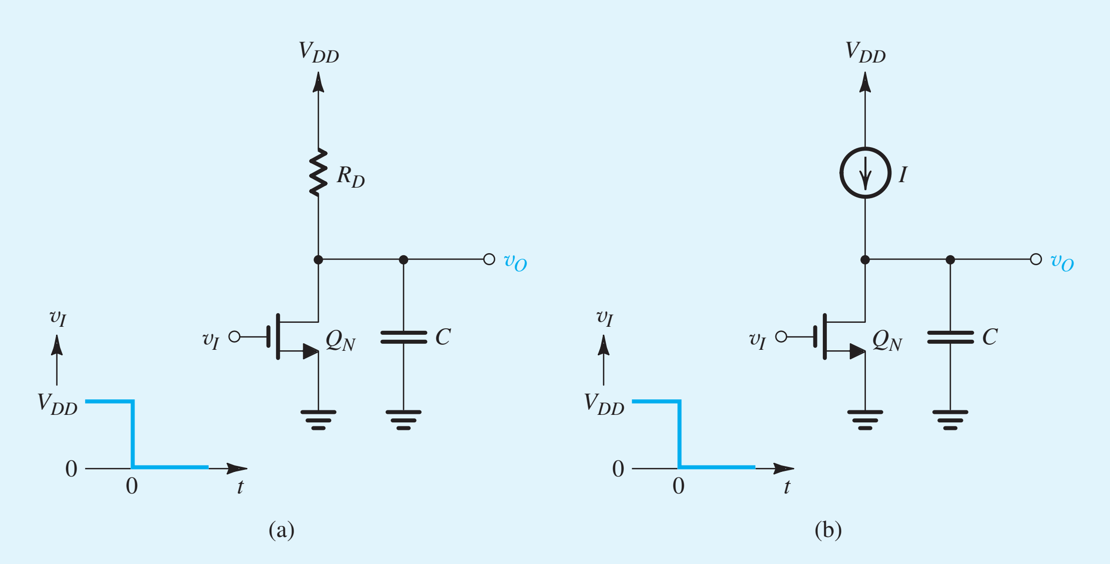


Figure P15.12

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

CHAPTER 15 PROBLEMS

D *15.20 It is required to design a minimum-area pseudo-NMOS inverter with equal high and low noise margins using a 1.3-V supply and devices for which $|V_t| = 0.4$ V, $k_n' = 4k_p' = 500\ \mu\text{A/V}^2$, and the minimum-size device has $(W/L) = 1$. Use $r = 5.7$ and show that $NM_L \simeq NM_H$. Specify the values of $(W/L)_n$ and $(W/L)_p$. What is the static power dissipated in this gate? What is the ratio of propagation delays for low-to-high and high-to-low transitions? For an equivalent load capacitance of 100 fF, find t_{PLH}, t_{PHL}, and t_P. At what frequency of operation would the static and dynamic power levels be equal? Is this speed of operation possible in view of the t_P value you found?

D 15.21 Sketch a pseudo-NMOS realization of the function $Y = \overline{A + B(C + D)}$.

D 15.22 Sketch a pseudo-NMOS realization of the exclusive-OR function $Y = A\overline{B} + \overline{A}B$.

D 15.23 Consider a four-input pseudo-NMOS NOR gate in which the NMOS devices have $(W/L)_n = 1.5$. It is required to find $(W/L)_p$ so that the worst-case value of V_{OL} is 0.1 V. Let $V_{DD} = 1.3$ V, $|V_t| = 0.4$ V, and $k_n' = 4k_p' = 500\ \mu\text{A/V}^2$.

15.24 This problem investigates the effect of velocity saturation (Section 15.1.3) on the operation of a pseudo-NMOS inverter fabricated in a 0.13-μm CMOS process for which $V_{DD} = 1.3$ V, $V_t = 0.4$ V, $k_n = 5k_p = 500\ \mu\text{A/V}^2$, and $|V_{DSsatp}| = 0.6$ V. Consider the case with $v_I = V_{DD}$ and $v_O = V_{OL}$. Note that Q_P will be operating in the velocity-saturation region. Find its current I_{Dsat} and use it to determine V_{OL}.

Section 15.4: Pass-Transistor-Logic Circuits

15.25 Recall that MOS transistors are symmetrical and that what distinguishes the source from the drain is their relative voltage levels: For NMOS, the terminal with the higher voltage is the drain; for PMOS, the terminal with the higher voltage is the source. For each of the circuits in Fig. P15.25, label the source and drain terminals and give the output voltage V_O in terms of V_{DD}, V_{tn}, and $|V_{tp}|$. Note that V_{tn} and $|V_{tp}|$ are determined by the body effect, and give expressions for their values. Note that V_O is the value reached after the capacitor charging/discharging interval has come to an end.

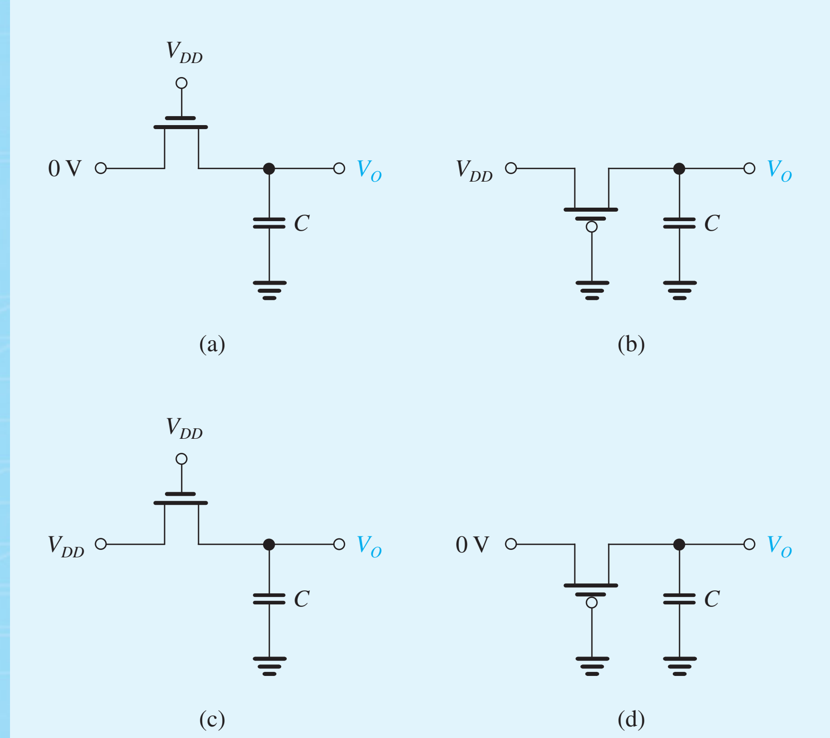


Figure P15.25

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

15.26 Let the NMOS transistor switch in Fig. 15.17 be fabricated in a 0.13-μm CMOS process for which $V_{t0} = 0.4$ V, $\gamma = 0.2\ \text{V}^{1/2}$, $2\phi_f = 0.88$ V, and $V_{DD} = 1.2$ V. Determine V_{OH}.

15.27 Consider the circuit in Fig. 15.17 with the NMOS transistor having $W/L = 1.5$ and fabricated in a CMOS process for which $V_{t0} = 0.4$ V, $\gamma = 0.2\ \text{V}^{1/2}$, $2\phi_f = 0.88$ V, $V_{DD} = 1.2$ V, and $\mu_n C_{ox} = 500\ \mu\text{A/V}^2$. Find t_{PLH} for the case $C = 10$ fF.

15.28 Consider the circuit in Fig. 15.18 with the NMOS transistor having $W/L = 1.5$ and fabricated in a 0.13-μm CMOS process for which $V_{t0} = 0.4$ V, $V_{DD} = 1.2$ V, and $\mu_n C_{ox} = 500\ \mu\text{A/V}^2$. Determine t_{PHL} for the case $C = 10$ fF.

15.29 Consider the case specified in Exercise 15.8. If the output of the switch is connected to the input of a CMOS inverter having $(W/L)_p = 2(W/L)_n = 0.54\ \mu\text{m}/0.18\ \mu\text{m}$, find the static current of the inverter and its static power dissipation when the inverter input is at the value found in Exercise 15.8. Also find the inverter output voltage. Let $\mu_n C_{ox} = 4\ \mu_p C_{ox} = 300\ \mu\text{A/V}^2$.

15.30 An NMOS pass-transistor switch with $W/L = 1.2\ \mu\text{m}/0.8\ \mu\text{m}$, used in a 3.3-V system for which $V_{t0} = 0.8$ V, $\gamma = 0.5\ \text{V}^{1/2}$, $2\phi_f = 0.6$V, $\mu_n C_{ox} = 3\mu_p C_{ox} = 75\ \mu\text{A/V}^2$, drives a 100-fF load capacitance at the input of a matched standard CMOS inverter using $(W/L)_n = 1.2\ \mu\text{m}/0.8\ \mu\text{m}$. For the switch gate terminal at V_{DD}, evaluate the switch V_{OH} and V_{OL} for inputs at V_{DD} and 0 V, respectively. For this value of V_{OH}, what inverter static current results? Estimate t_{PLH} and t_{PHL} for this arrangement as measured from the input to the output of the switch itself.

15.31 Figure P15.31 shows a PMOS transistor operating as a switch in the on position.

(a) If initially $v_O = 0$ and at $t = 0$, v_I is raised to V_{DD}, what is the final value V_{OH} reached at the output?

(b) If initially $v_O = V_{DD}$ and at $t = 0$, v_I is lowered to 0 V, what is the final value V_{OL} reached at the output?

(c) For the situation in (a), find t_{PLH} for v_O to rise from 0 to $V_{DD}/2$. Let $k_p = 125\ \mu\text{A/V}^2$, $V_{DD} = 1.2$ V, and $|V_{tp}| = 0.4$ V.

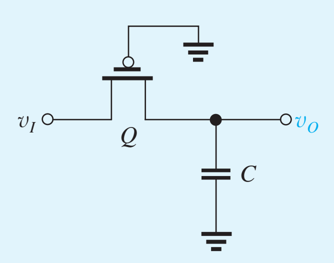


Figure P15.31

***15.32** A designer, beginning to experiment with the idea of pass-transistor logic, seizes upon what he sees as two good ideas:

(a) that a string of minimum-size single MOS transistors can do complex logic functions,

(b) but that there must always be a path between output and a supply terminal.

Correspondingly, he first considers two circuits (shown in Fig. P15.32). For each, express Y as a function of A and B. In each case, what can be said about general operation? About the logic levels at Y? About node X? Do either of these circuits look familiar? If in each case the terminal connected to V_{DD} is instead connected to the output of a CMOS inverter whose input is connected to a signal C, what does the function Y become?

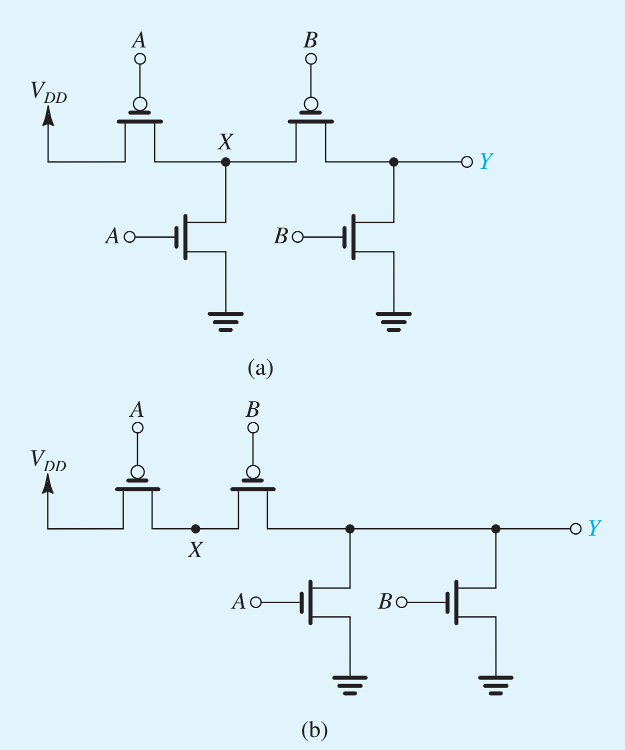


Figure P15.32

15.33 Consider the circuits in Fig. P15.32 with all PMOS transistors replaced with NMOS, and all NMOS by PMOS, and with ground and V_{DD} connections interchanged. What do the output functions Y become?

15.34 For the level-restoring circuit of Fig. 15.19, let $k_n' = 3k_p' = 75\ \mu A/V^2$, $V_{DD} = 3.3$ V, $|V_{t0}| = 0.8$ V, $\gamma = 0.5\ V^{1/2}$, $2\phi_f = 0.6$ V, $(W/L)_1 = (W/L)_n = 1.2\ \mu m/0.8\ \mu m$, $(W/L)_p = 3.6\ \mu m/0.8\ \mu m$, and $C = 20$ fF. Also let $v_B = V_{DD}$. Now for v_A rising to V_{DD}, and Q_1 charging C and causing v_{O1} to rise, show that the value of v_{O1} that causes v_{O2} to drop by a threshold voltage below V_{DD} (i.e., to 2.5 V) so that Q_R turns on, is approximately $V_{DD}/2$ and thus occurs at $t \simeq t_{LH}$. What is the capacitor-charging current available at this time (i.e., just prior to Q_R turning on)? What is it at $v_{O1} = 0$? What is the average current available for charging C? Estimate the time t_{PLH}. (Note that after Q_R turns on, v_{O1} rises to V_{DD}.)

D *15.35 The purpose of this problem is to illustrate how W/L of the level-restoring transistor Q_R in the circuit of Fig. 15.19 is determined. For this purpose consider the circuit as specified in Problem 15.34 and let $v_B = V_{DD}$. Now, consider the situation when v_A is brought down to 0 V and Q_1 conducts and begins to discharge C. The voltage v_{O1} will begin to drop from V_{DD}. Meanwhile, v_{O2} is still low and Q_R is conducting (though at $t = 0$, the current in Q_R is zero). Calculate the discharge current at $t = 0$. As Q_R conducts, its current subtracts from the current of Q_1, reducing the current available to discharge C. Find the value of v_{O1} at which the inverter begins to switch. This is $V_{IH} = \frac{1}{8}(5V_{DD} - 2V_t)$. Then, find the current that Q_1 conducts at this value of v_{O1}. Choose W/L for Q_R so that the maximum current it conducts is limited to one-half the value of the current in Q_1. What is the W/L you have chosen? Estimate t_{PHL} as the time for v_{O1} to drop from V_{DD} to V_{IH}.

15.36 The transmission gate in Fig. 15.21(a) and 15.21(b) is fabricated in a CMOS process technology for which $k_n' = 4k_p' = 500\ \mu A/V^2$, $|V_{t0}| = 0.4$ V, $\gamma = 0.2\ V^{1/2}$, $2\phi_f = 0.88$ V, and $V_{DD} = 1.2$ V. Let Q_N and Q_P have $(W/L)_n = (W/L)_p = 1.5$. The total capacitance at the output node is 15 fF.

(a) What are the values of V_{OH} and V_{OL}?
(b) For the situation in Fig. 15.21(a), find $i_{DN}(0)$, $i_{DP}(0)$, $i_{DN}(t_{PLH})$, $i_{DP}(t_{PLH})$, and t_{PLH}.
(c) For the situation depicted in Fig. 15.21(b), find $i_{DN}(0)$, $i_{DP}(0)$, $i_{DN}(t_{PHL})$, $i_{DP}(t_{PHL})$, and t_{PHL}. At what value of v_O will Q_P turn off?
(d) Find t_P.

15.37 For the transmission gate specified in Problem 15.36, find R_{TG} at $v_O = 0$ and 0.6 V. Use the average of those values to determine t_{PLH} for the situation in which $C = 15$ fF.

***15.38** Refer to the situation in Fig. 15.21(b). Derive expressions for R_{Neq}, R_{Peq}, and R_{TG} following the approach used in Section 15.4.4 for the capacitor-charging case. Evaluate the value of R_{TG} for $v_O = V_{DD}$ and $v_O = V_{DD}/2$ for the process technology specified in Problem 15.36. Find the average value of R_{TG} and use it to determine t_{PHL} for the case $C = 15$ fF.

15.39 A transmission gate for which $(W/L)n = (W/L)_p = 1.5$ is fabricated in a 0.13-μm CMOS technology and used in a circuit for which $C = 10$ fF. Use Eq. (15.49) to obtain an estimate of R_{TG} and hence of the propagation delay t_P.

15.40 Figure P15.40 shows a chain of transmission gates. This situation often occurs in circuits such as adders and multiplexers. Consider the case when all the transmission gates are turned on and a step voltage V_{DD} is applied to the input. The propagation delay t_P can be determined from the Elmore delay formula as follows:

$$t_P = 0.69 \sum_{k=0}^{n} kCR_{TG}$$

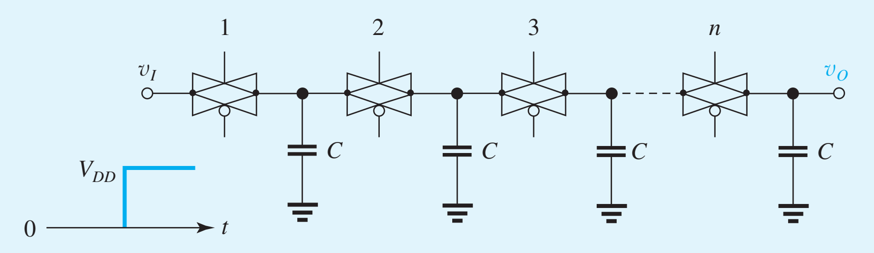


Figure P15.40

where R_{TG} is the resistance of each transmission gate, C is the capacitance between each node and ground, and n is the number of transmission gates in the chain. Note that the sum of the series in this formula is given by

$$t_P = 0.69\, CR_{TG}\frac{n(n+1)}{2}$$

Now evaluate t_P for the case of 16 transmission gates with $R_{TG} = 10\,\text{k}\Omega$ and $C = 10$ fF. What does the value of t_P become if the input is a ramp rather than a step function?

D 15.41 (a) Use the idea embodied in the exclusive-OR realization in Fig. 15.26 to realize $\overline{Y} = AB + \overline{A}\,\overline{B}$. That is, find a realization for $\overline{Y}$ using two transmission gates.
(b) Now combine the circuit obtained in (a) with the circuit in Fig. 15.26 to obtain a realization of the function $Z = \overline{Y}C + Y\overline{C}$ where C is a third input. Sketch the complete 12-transistor circuit realization of Z. Note that Z is a three-input exclusive-OR.

D *15.42 Using the idea presented in Fig. 15.27, sketch a CPL circuit whose outputs are $Y = A\overline{B} + \overline{A}B$ and $\overline{Y} = AB + \overline{A}\,\overline{B}$.

D 15.43 Extend the CPL idea in Fig. 15.27 to three variables to form $Z = ABC$ and $\overline{Z} = \overline{ABC} = \overline{A} + \overline{B} + \overline{C}$.

Section 15.5: Dynamic MOS Logic Circuits

D 15.44 Based on the basic dynamic logic circuit of Fig. 15.28, sketch complete circuits for NOT, NAND, and NOR gates, the latter two with two inputs, and a circuit for which $Y = \overline{AB + CD}$.

15.45 In this and the following problem, we investigate the dynamic operation of a two-input NAND gate realized in the dynamic logic form and fabricated in a CMOS process technology for which $k_n' = 4k_p' = 500\ \mu\text{A/V}^2$, $V_{tn} = -V_{tp} = 0.4$ V, and $V_{DD} = 1.2$ V. To keep C_L small, minimum-size NMOS devices are used for which $W/L = 1.5$ (this includes Q_e). The PMOS precharge transistor Q_p has $W/L = 3$. The capacitance C_L is found to be 15 fF. Consider the precharge operation with the gate of Q_p at 0 V, and assume that at $t = 0$, C_L is fully discharged. Calculate the rise time of the output voltage, defined as the time for v_Y to rise from 10% to 90% of the final value of 1.2 V.

15.46 For the gate specified in Problem 15.45, evaluate the high-to-low propagation delay, t_{PHL}.

15.47 Consider a two-input NOR gate realized in the dynamic logic form illustrated in Fig. 15.28. Assume that the gate is fabricated in a 0.13-μm CMOS technology for which $V_{DD} = 1.2$ V, $V_t = 0.4$ V, and $\mu_n C_{ox} = 4\,\mu_p C_{ox} = 500\ \mu\text{A/V}^2$. The NMOS devices have $W/L = 1.5$, and the PMOS transistor has $W/L = 3$. The total capacitance at the output is found to be 15 fF. Calculate the rise time of v_O, from 0.1 V_{DD} to 0.9 V_{DD}, in the precharge interval. Also, calculate the worst-case value of t_{PHL}.

15.48 The leakage current in a dynamic logic gate causes the capacitor C_L to discharge during the evaluation phase, even if the PDN is not conducting. For $C_L = 10$ fF, and $I_{\text{leakage}} = 2 \times 10^{-12}$ A, find the longest allowable evaluation time if the decay in output voltage is to be limited to 0.2 V. If the precharge interval is much shorter than the maximum allowable evaluation time, find the minimum clocking frequency required.

***15.49** In this problem, we wish to calculate the reduction in the output voltage of a dynamic logic gate as a result of charge redistribution. Refer to the circuit in Fig. 15.30(a), and assume that at $t = 0-$, $v_Y = V_{DD}$, and $v_{C1} = 0$. At $t = 0$, ϕ goes high and Q_P turns off, and simultaneously the voltage at the gate of Q_1 goes high (to V_{DD}), turning Q_1 on. Transistor Q_1 will remain conducting until either the voltage at its source (v_{C1}) reaches $V_{DD} - V_{tn}$ or until $v_Y = v_{C1}$, whichever comes first. In both cases, the final value of v_Y can be found using charge conservation: that is, by equating the charge gained by C_1 to the charge lost by C_L.

(a) Convince yourself that the first situation obtains when $|\Delta v_Y| \le V_{tn}$.
(b) For each of the two situations, derive an expression for Δv_Y.
(c) Find an expression for the maximum ratio (C_1/C_L) for which $|\Delta v_Y| \le V_{tn}$.
(d) For $V_{tn} = 0.5$ V, $V_{DD} = 1.8$ V, $C_L = 15$ fF, and neglecting the body effect in Q_1, find the drop in voltage at the output in the two cases: (a) $C_1 = 4$ fF and (b) $C_1 = 7.5$ fF.

15.50 Solve the problem in Exercise 15.15 symbolically (rather than numerically). Refer to Fig. E15.15 and assume Q_{eq1} and Q_{eq2} to be identical with threshold voltages $V_{tn} = 0.2V_{DD}$ and transconductance parameters k_n. Also, let $C_{L1} = C_{L2}$. Derive an expression for the drop in the output voltage, Δv_{Y2}.

15.51 For the four-input dynamic logic NAND gate analyzed in Example 15.4, estimate the maximum clocking frequency allowed.

Section 15.6: Bipolar and BiCMOS Logic Circuits

15.52 For the circuit in Fig. 15.34, let $V_{CC} = 0$ V, $I = 1$ mA, and $V_R = -1$ V. Find R_C to obtain an output voltage swing of 0.4 V. By how much should the output levels be shifted so that the values of V_{OH} and V_{OL} become centered on V_R? What will the shifted values of V_{OH} and V_{OL} be?

15.53 In Fig. P15.53(a), the ECL gate discussed in the text, only the input A is shown (the other input B is assumed to be left open and thus deactivated). Figure P15.53(b) shows that OR transfer characteristic: that is, v_{OR} versus v_I. Determine the parameters of the transfer characteristic: that is, V_{IL}, V_{IH}, V_{OL}, and V_{OH}. Define V_{IL} as the value of v_I for which Q_R conducts 99% of I_E and Q_A conducts 1% of I_E. Conversely, define V_{IH} as the value of v_I for which Q_A conducts 99% of I_E and Q_R conducts 1% of I_E. Also, determine the width of the transition region (i.e., $V_{IH} - V_{IL}$) and the noise margins NM_H and NM_L. Assume that at an emitter current of 1 mA the transistor $V_{BE} = 0.75$ V and $\beta_2 = 100$.

D 15.54 For the ECL circuit in Fig. P15.54, the transistors exhibit V_{BE} of 0.75 V at an emitter current I and have very high β.

(a) Find V_{OH} and V_{OL}.

(b) For the input at B that is sufficiently negative for Q_B to be cut off, what voltage at A causes a current of $I/2$ to flow in Q_R?

(c) Repeat (b) for a current in Q_R of 0.99I. Define this value of v_A as V_{IL}.

(d) Repeat (c) for a current in Q_R of 0.01I. Define this value of v_A as V_{IH}.

(e) Use the results of (c) and (d) to specify V_{IL} and V_{IH}.

(f) Find NM_H and NM_L.

(g) Find the value of IR that makes the noise margins equal to the width of the transition region, $V_{IH} - V_{IL}$.

(h) Using the IR value obtained in (g), give numerical values for V_{OH}, V_{OL}, V_{IH}, V_{IL}, and V_R for this ECL gate.

15.55 For the ECL gate in Fig. 15.35, calculate an approximate value for the power dissipated in the circuit under the conditions that all inputs are low and that the emitters of the output followers are left open. Assume that the reference circuit supplies four identical gates, and hence only a quarter of the power dissipated in the reference circuit should be attributed to a single gate.

D *15.56 Using the logic and circuit flexibility of ECL indicated by Figs. 15.35 and 15.37, sketch an ECL logic circuit that realizes the exclusive-OR function, $Y = \bar{A}B + A\bar{B}$. Give a logic diagram (as opposed to a circuit diagram).

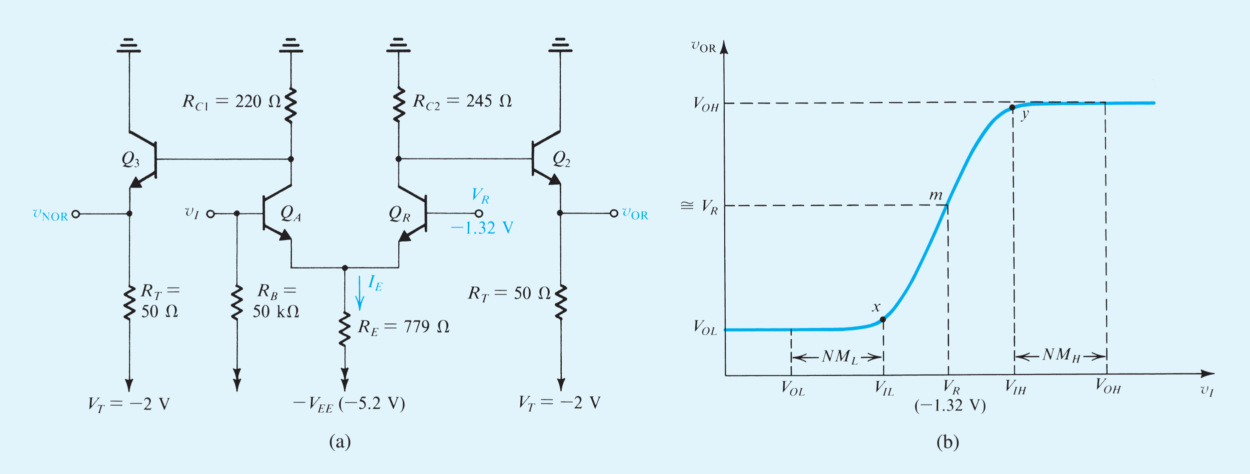


Figure P15.53

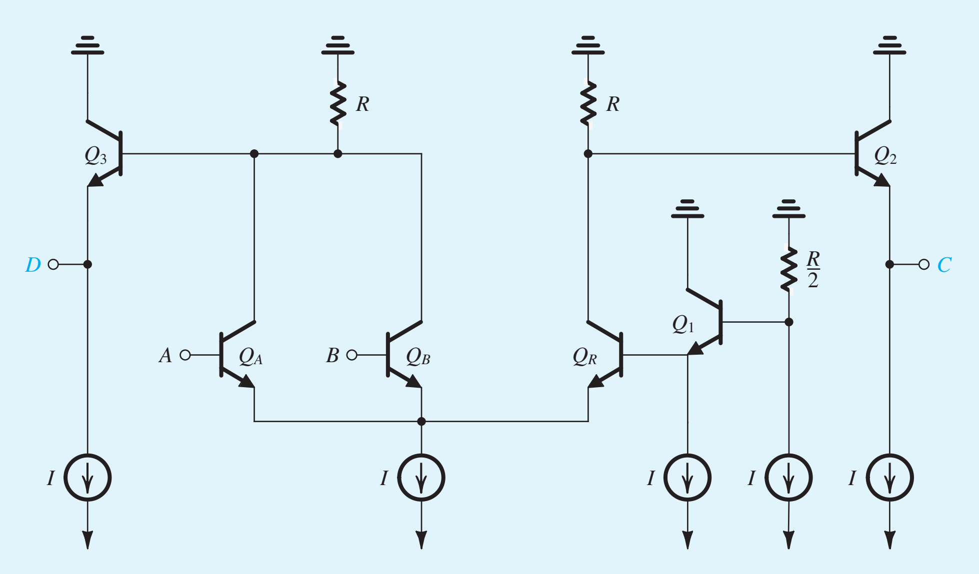


Figure P15.54

***15.57** For the circuit in Fig. P15.57, let the levels of the inputs A, B, C, and D be 0 and +5 V. For all inputs low at 0 V, what is the voltage at E? If A and C are raised to +5 V, what is the voltage at E? Assume $|V_{BE}| = 0.7$ V and $\beta = 50$. Express E as a logic function of A, B, C, and D.

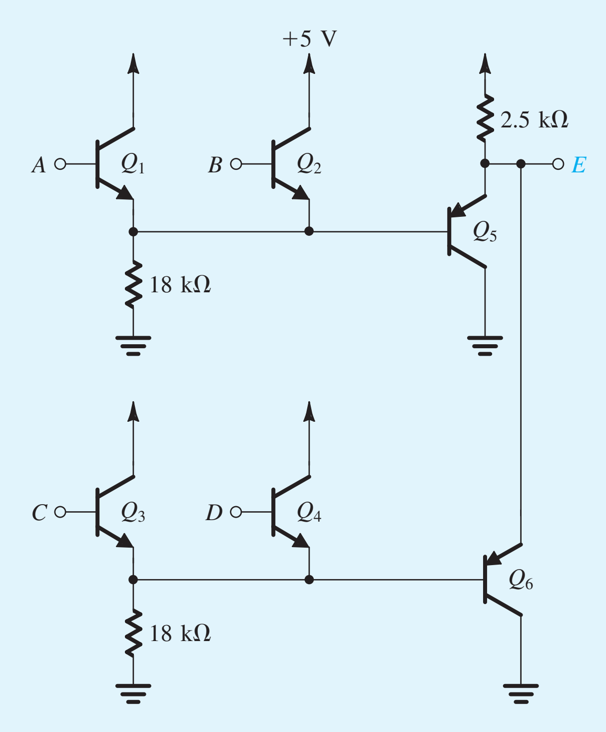


Figure P15.57

15.58 Consider the conceptual BiCMOS circuit of Fig. 15.38(a), for the conditions that $V_{DD} = 5$ V, $|V_t| = 1$ V, $V_{BE} = 0.7$ V, $\beta = 100$, $k_n' = 2.5k_p' = 100\,\mu\text{A/V}^2$, and $(W/L)_n = 2\ \mu\text{m}/1\,\mu\text{m}$. For $v_I = v_O = V_{DD}/2$, find $(W/L)_p$ so that $I_{EQ1} = I_{EQ2}$. What is this totem-pole transient current?

15.59 Consider the conceptual BiCMOS circuit of Fig. 15.38(a) for the conditions stated in Problem 15.58. What is the threshold voltage of the inverter if both Q_N and Q_P have $W/L = 2\ \mu\text{m}/1\ \mu\text{m}$? What totem-pole current flows at v_I equal to the threshold voltage?

D *15.60 Consider the choice of values for R_1 and R_2 in the circuit of Fig. 15.38(c). Let the inverter be specified as in Problem 15.58 with the MOSFETs matched and $(W/L)_P = 2.5(W/L)_n$. An important consideration in making this choice is that the loss of base drive current will be limited. This loss becomes particularly acute when the current through Q_N and Q_P becomes small. This in turn happens near the end of the output signal swing when the associated MOS device is deeply in triode operation (say at $|v_{DS}| = |V_t|/3$). Determine values for R_1 and R_2 so that the loss in base current is limited to 50%. What is the ratio R_1/R_2? Repeat for a 20% loss in base drive.

D 15.61 Sketch the circuit of a BiCMOS two-input NOR gate based on the R-circuit of Fig. 15.38(e).

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

CHAPTER 16

Memory Circuits

IN THIS CHAPTER YOU WILL LEARN

1. How the basic bistable circuit, the latch, is realized by connecting two inverters in a positive-feedback loop.
2. How to augment the latch to obtain different types of flip-flops that are useful building blocks for digital systems.
3. How CMOS is particularly suited for the efficient implementation of a particular type of flip-flop, the *D* flip-flop.
4. How memory chips that contain as many as 4 gigabits are organized, as well as their various types and the terminology used to describe them.
5. The analysis and design of the six-transistor circuit that is almost universally used to implement the storage cell in static random-access memory (SRAM) and the one-transistor circuit that is equally universal in the implementation of the storage cell in dynamic random-access memory (DRAM).
6. Interesting circuit techniques for accessing a particular storage cell in a memory chip and for amplifying the signal readout from the cell.
7. How various types of read-only memory (ROM) are designed, programmed, erased, and reprogrammed.
8. How the basic image-capturing element works in digital cameras (including smartphone cameras).

Introduction

The logic circuits studied in Chapters 14 and 15 are called **combinational circuits**. Their output depends only on the present value of the input. Thus these circuits do *not* have memory. *Memory* is a very important part of digital systems. Its availability in digital computers allows for storing programs and data. Furthermore, it is important for temporary storage of the output produced by a combinational circuit for use at a later time in the operation of a digital system.

Logic circuits that incorporate memory are called **sequential circuits**; that is, their output depends not only on the present value of the input but also on the input's previous values. Such circuits require a timing generator (a *clock*) for their operation.

There are basically two approaches for providing memory to a digital circuit. The first relies on the application of positive feedback that, as will be seen shortly, can be arranged

to provide a circuit with two stable states. Such a *bistable* circuit can then be used to store one bit of information: One stable state would correspond to a stored 0, and the other to a stored 1. A bistable circuit can remain in either state indefinitely, and thus it belongs to the category of *static sequential circuits*. The other approach to realizing memory utilizes the storage of charge on a capacitor: When the capacitor is charged, it would be regarded as storing a 1; when it is discharged, it would be storing a 0. Since the inevitable leakage effects will cause the capacitor to discharge, such a form of memory requires the periodic recharging of the capacitor, a process known as *refresh*. Thus, like dynamic logic (Section 15.5), memory based on charge storage is known as *dynamic memory* and the corresponding sequential circuits as *dynamic sequential circuits*.

This chapter is concerned with the study of memory circuits. We begin in Section 16.1 with the basic bistable circuit, the latch, and its application in flip-flops, an important class of building blocks for digital systems. After an overview of memory-chip types, organization, and nomenclature in Section 16.2, we study the circuit of the static memory cell (SRAM) and that of the dynamic memory cell (DRAM) in Section 16.3. Besides the array of storage cells, memory chips require circuits for selecting and accessing a particular cell in the array (address decoders) and for amplifying the signal that is retrieved from a particular cell (sense amplifiers). A sampling of these peripheral circuits is presented in Section 16.4. An important class of memories, the read-only memory (ROM) is presented in Section 16.5. The chapter concludes with a brief look at a very important system component, the CMOS image sensor. Having a structure very similar to that of a memory array, the image sensor is the basic image-capturing element in digital cameras (including smartphone cameras).

16.1 Latches and Flip-Flops

In this section, we shall study the basic memory element, the latch, and consider a sampling of its applications. Both static and dynamic circuits will be considered.

16.1.1 The Latch

The basic memory element, the latch, is shown in Fig. 16.1(a). It consists of two cross-coupled logic inverters, G_1 and G_2. The inverters form a positive-feedback loop. To investigate the operation of the latch we break the feedback loop at the input of one of the inverters, say G_1, and apply an input signal, v_W, as shown in Fig. 16.1(b). Assuming that the input impedance of G_1 is large, breaking the feedback loop will not change the loop voltage-transfer characteristic, which can be determined from the circuit of Fig. 16.1(b) by plotting v_Z versus v_W. This is the voltage-transfer characteristic of two cascaded inverters and thus takes the shape shown in Fig. 16.1(c). Observe that the transfer characteristic consists of three segments, with the middle segment corresponding to the transition region of the inverters.

Also shown in Fig. 16.1(c) is a straight line with unity slope. This straight line represents the relationship $v_W = v_Z$ that is realized by reconnecting Z to W to close the feedback loop and thus to return it to its original form. As indicated, the straight line intersects the loop transfer curve at three points, A, B, and C. Thus any of these three points can serve as the operating point for the latch. We shall now show that while points A and C are stable operating points in the sense that the circuit can remain at either one indefinitely, point B is an unstable operating point; the latch cannot operate at B for any significant period of time.

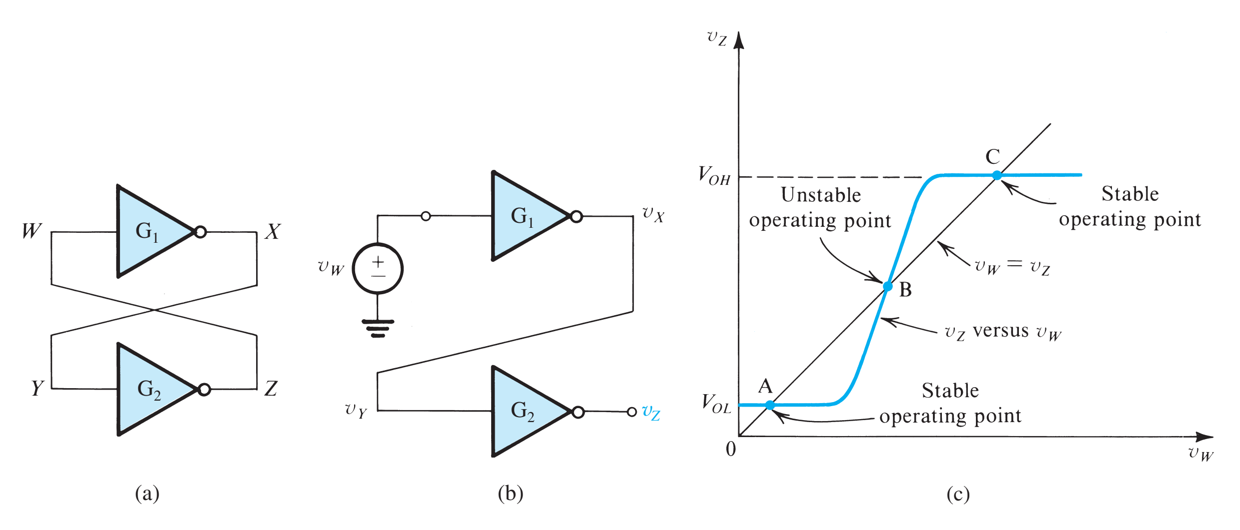


Figure 16.1 **(a)** Basic latch. **(b)** The latch with the feedback loop opened. **(c)** Determining the operating point(s) of the latch.

The reason point B is unstable can be seen by considering the latch circuit in Fig. 16.1(a) to be operating at point B, and taking account of the electrical interference (or noise) that is inevitably present in any circuit. Let the voltage v_W increase by a small increment v_w. The voltage at *X* will increase (in magnitude) by a larger increment, equal to the product of v_w and the incremental gain of G_1 at point B. The resulting signal v_x is applied to G_2 and gives rise to an even larger signal at node *Z*. The voltage v_z is related to the original increment v_w by the loop gain at point B, which is the slope of the curve of v_Z versus v_W at point B. This gain is usually much greater than unity. Since v_z is coupled to the input of G_1, it becomes the new value of v_W and is further amplified by the loop gain. This regenerative process continues, shifting the operating point from B upward to point C, as illustrated in Fig. 16.2. Since at C the loop gain is zero (or almost zero), no regeneration can take place.

In the description above, we assumed arbitrarily an initial positive voltage increment at *W*. Had we instead assumed a negative voltage increment, we would have seen that the operating point moves downward from B to A. Again, since at point A the slope of the transfer curve is zero (or almost zero), no regeneration can take place. In fact, for regeneration to occur, the loop gain must be greater than unity, which is the case at point B.

The discussion above leads us to conclude that the latch has two stable operating points, A and C. At point C, v_W is high, v_X is low, v_Y is low, and v_Z is high. The reverse is true at point A. If we consider *X* and *Z* as the latch outputs, we see that in one of the stable states (say that corresponding to operating point A), v_X is high (at V_{OH}) and v_Z is low (at V_{OL}). In the other state (corresponding to operating point C), v_X is low (at V_{OL}) and v_Z is high (at V_{OH}). Thus the latch is a **bistable** circuit having two complementary outputs. The stable state in which the latch operates depends on the external excitation that forces it to the particular state. The latch then *memorizes* this external action by staying indefinitely in the acquired state. As a memory element the latch is capable of storing one bit of information. For instance, we can arbitrarily designate the state in which v_X is high and v_Z is low as corresponding to a stored logic 1. The other complementary state then is designated by a stored logic 0. Finally, we note that the latch circuit described is of the static variety.

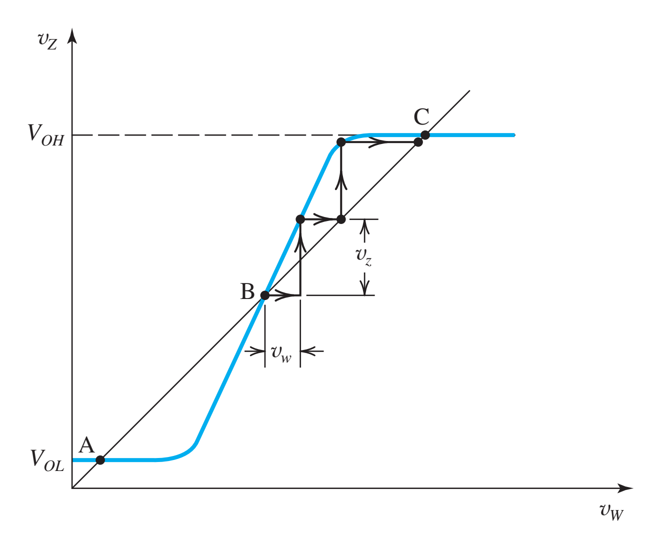


Figure 16.2 Point B is an unstable operating point for the latch: A small positive increment v_w gets amplified around the loop and causes the operating point to shift to the stable operating point C. Had v_w been negative, the operating point would have shifted to the other stable point, A.

It now remains to devise a mechanism by which the latch can be *triggered* to change state. The latch together with the triggering circuitry forms a *flip-flop*. This will be discussed next. Analog bistable circuits utilizing op amps will be presented in Chapter 18.

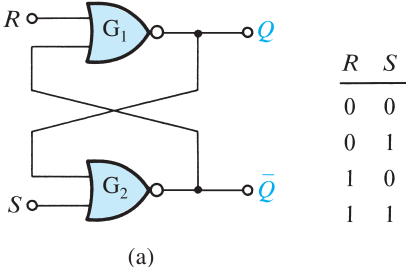
16.1.2 The SR Flip-Flop

The simplest type of flip-flop is the set/reset (SR) flip-flop shown in Fig. 16.3(a). It is formed by cross-coupling two NOR gates, and thus it incorporates a latch. The second inputs of G_1 and G_2 together serve as the trigger inputs of the flip-flop. These two inputs are labeled S (for set) and R (for reset). The outputs are labeled Q and $\overline{Q}$, emphasizing their complementarity. The flip-flop is considered to be set (i.e., storing a logic 1) when Q is high and $\overline{Q}$ is low. When the flip-flop is in the other state (Q low, $\overline{Q}$ high), it is considered to be reset (storing a logic 0).

In the *rest* or *memory state* (i.e., when we do not wish to change the state of the flip-flop), both the S and R inputs should be low. Consider the case when the flip-flop is storing a logic 0. Since Q will be low, both inputs to the NOR gate G_2 will be low. Its output will therefore be high. This high is applied to the input of G_1, causing its output Q to be low, satisfying the original assumption. To set the flip-flop we raise S to the logic-1 level while leaving R at 0.

FLIP-FLOP FACT:

In 1918 William H. Eccles and Frank W. Jordan, while working at City and Guilds Technical College in London, filed the first patent for an electronic bistable element, the bellwether of the digital electronic age. The initial design captured the idea of a closed positive-feedback loop involving two inverting direct-coupled amplifiers, along with “gating” to allow signals to be injected. While the first designs used vacuum tubes, the same structure was extended to bipolar and MOS transistors as those technologies became commercially available. The Eccles–Jordan flip-flop remains, nearly a century later, as an enormous milestone in digital computing.



R	S	Q_{n+1}
0	0	Q_n
0	1	1
1	0	0
1	1	Not used

(b)

Figure 16.3 **(a)** The set/reset (SR) flip-flop and **(b)** its truth table.

The 1 at the S terminal will force the output of G_2, $\overline{Q}$, to 0. Thus the two inputs to G_1 will be 0 and its output Q will go to 1. Now even if S returns to 0, the $Q = 1$ signal fed to the input of G_2 will keep $\overline{Q} = 0$, and the flip-flop will remain in the newly acquired set state. Note that if we raise S to 1 again (with R remaining at 0), no change will occur. To reset the flip-flop we need to raise R to 1 while leaving $S = 0$. We can readily show that this forces the flip-flop into the reset state ($Q = 0, \overline{Q} = 1$) and that the flip-flop remains in this state even after R has returned to 0. It should be observed that the trigger signal merely starts the regenerative action of the positive-feedback loop of the latch.

Finally, we inquire into what happens if both S and R are simultaneously raised to 1. The two NOR gates will cause both Q and $\overline{Q}$ to become 0 (note that in this case the complementary labeling of these two variables is incorrect). However, if R and S return to the rest state ($R = S = 0$) simultaneously, the state of the flip-flop will be undefined. In other words, it will be impossible to predict the final state of the flip-flop. For this reason, this input combination is usually disallowed (i.e., not used). Note, however, that this situation arises only in the idealized case, when both R and S return to 0 precisely simultaneously. In actual practice one of the two will return to 0 first, and the final state will be determined by the input that remains high longest.

The operation of the flip-flop is summarized by the *truth table* in Fig. 16.3(b), where Q_n denotes the value of Q at time t_n just before the application of the R and S signals, and Q_{n+1} denotes the value of Q at time t_{n+1} after the application of the input signals.

Rather than using two NOR gates, one can also implement an SR flip-flop by cross-coupling two NAND gates, in which case the set and reset functions are active when low (see Problem 16.3).

16.1.3 CMOS Implementation of SR Flip-Flops

The SR flip-flop of Fig. 16.3 can be directly implemented in CMOS by simply replacing each of the NOR gates by its CMOS circuit realization. We encourage the reader to sketch the resulting circuit (see Problem 16.2). Although the CMOS circuit thus obtained works well, it is somewhat complex. As an alternative, we consider a simplified circuit that furthermore implements additional logic. Specifically, Fig. 16.4 shows a *clocked* version of an SR flip-flop. Since the clock inputs form AND functions with the set and reset inputs, the flip-flop can be set or reset only when the clock ϕ is high. Observe that although the two cross-coupled inverters at the heart of the flip-flop are of the standard CMOS type, only NMOS transistors are used for the set–reset circuitry. Nevertheless, since there is no conducting path between V_{DD} and ground (except during switching), the circuit does not dissipate any static power.

Except for the addition of clocking, the SR flip-flop of Fig. 16.4 operates in exactly the same way as its logic antecedent in Fig. 16.3: To illustrate, consider what happens when the

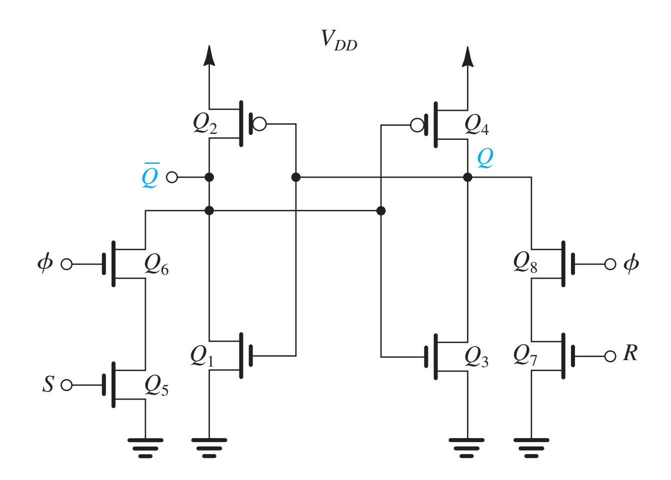


Figure 16.4 CMOS implementation of a clocked SR flip-flop. The clock signal is denoted by ϕ.

flip-flop is in the reset state ($Q = 0, \overline{Q} = 1$, $v_Q = 0$, $v_{\overline{Q}} = V_{DD}$), and assume that we wish to set it. To do so, we arrange for a high (V_{DD}) signal to appear on the S input while R is held low at 0 V. Then, when the clock ϕ goes high, both Q_5 and Q_6 will conduct, pulling the voltage $v_{\overline{Q}}$ down. If $v_{\overline{Q}}$ goes below the threshold V_M of the (Q_3, Q_4) inverter, the inverter will switch states (or at least begin to switch states), and its output v_Q will rise. This increase in v_Q is fed back to the input of the (Q_1, Q_2) inverter, causing its output $v_{\overline{Q}}$ to go down even further; the regeneration process, characteristic of the positive-feedback latch, is now in progress.

The preceding description of flip-flop switching is predicated on two assumptions:

1. Transistors Q_5 and Q_6 supply sufficient current to pull the node $\overline{Q}$ down to a voltage at least slightly below the threshold of the (Q_3, Q_4) inverter. This is essential for the regenerative process to begin. Without this initial trigger, the flip-flop will fail to switch. In Example 16.1, we shall investigate the *minimum W/L* ratios that Q_5 and Q_6 must have to meet this requirement.
2. The set signal remains high for an interval long enough to cause regeneration to take over the switching process. An estimate of the minimum width required for the set pulse can be obtained as the sum of the interval during which $v_{\overline{Q}}$ is reduced from V_{DD} to $V_{DD}/2$, and the interval for the voltage v_Q to respond and rise to $V_{DD}/2$. This point also will be illustrated in Example 16.1.

Finally, note that the symmetry of the circuit indicates that all the preceding remarks apply equally well to the reset process.

Example 16.1

The CMOS SR flip-flop in Fig. 16.4 is fabricated in a 0.18-μm process for which $\mu_n C_{ox} = 4\mu_p C_{ox} = 300\ \mu\text{A/V}^2$, $V_{tn} = |V_{tp}| = 0.5$ V, and $V_{DD} = 1.8$ V. The inverters have $(W/L)_n = 0.27\ \mu\text{m}/0.18\ \mu\text{m}$ and $(W/L)_p = 4(W/L)_n$. The four NMOS transistors in the set–reset circuit have equal W/L ratios.

(a) Determine the minimum value required for this ratio to ensure that the flip-flop will switch.
(b) Also, determine the minimum width the set pulse must have for the case in which the W/L ratio of each of the four transistors in the set–reset circuit is selected at twice the minimum value found in (a). Assume that the total capacitance between each of the Q and $\overline{Q}$ nodes and ground is 20 fF.

Solution

(a) Figure 16.5(a) shows the relevant portion of the circuit for our present purposes. Observe that since the circuit is in the reset state and regeneration has not yet begun, we assume that $v_Q = 0$ and thus Q_2 will be conducting. The circuit is in effect a pseudo-NMOS gate, and our task is to select the W/L ratios for Q_5 and Q_6 so that V_{OL} of this inverter is lower than $V_{DD}/2$ (the threshold of the Q_3, Q_4 inverter whose Q_N and Q_P are matched). The minimum required W/L for Q_5 and Q_6 can be found by equating the current supplied by Q_5 and Q_6 to the current supplied by Q_2 at $v_{\overline{Q}} = V_{DD}/2$. To simplify matters, we assume that the series connection of Q_5 and Q_6 is equivalent to a single transistor whose W/L is half the W/L of each of Q_5 and Q_6 [Fig. 16.5(b)]. Now, since at $v_{\overline{Q}} = V_{DD}/2 = 0.9$ V and $|V_t| = 0.5$ V, both this equivalent transistor and Q_2 will be operating in the triode region, we can write

$$I_{Deq} = I_{D2}$$

$$300 \times \frac{1}{2}\left(\frac{W}{L}\right)_5 \left[(1.8 - 0.5)\left(\frac{1.8}{2}\right) - \frac{1}{2}\left(\frac{1.8}{2}\right)^2\right]$$

$$= 75 \times \frac{1.08}{0.18}\left[(1.8 - 0.5)\left(\frac{1.8}{2}\right) - \frac{1}{2}\left(\frac{1.8}{2}\right)^2\right]$$

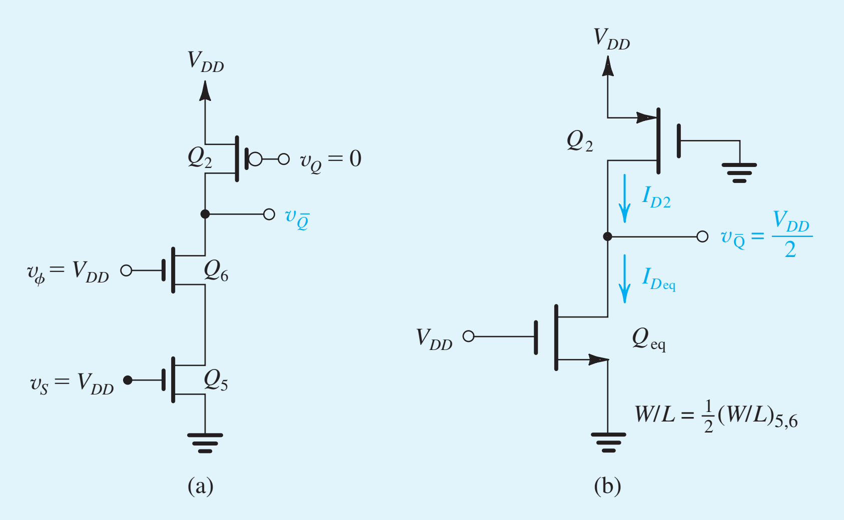


Figure 16.5 **(a)** The relevant portion of the flip-flop circuit of Fig. 16.4 for determining the minimum W/L ratios of Q_5 and Q_6 needed to ensure that the flip-flop will switch. **(b)** The circuit in **(a)** with Q_5 and Q_6 replaced with their equivalent transistor Q_{eq}, at the point of switching.

Example 16.1 *continued*

which yields

$$\left(\frac{W}{L}\right)_5 = \frac{0.54\ \mu\text{m}}{0.18\ \mu\text{m}}$$

and thus

$$\left(\frac{W}{L}\right)_6 = \frac{0.54\ \mu\text{m}}{0.18\ \mu\text{m}}$$

(b) The value calculated for $(W/L)_5$ and $(W/L)_6$ is the absolute minimum needed for switching to occur. To guarantee that the flip-flop will switch, the value selected for $(W/L)_5$ and $(W/L)_6$ is usually somewhat larger than the minimum. Selecting a value twice the minimum,

$$(W/L)_5 = (W/L)_6 = 1.08\ \mu\text{m}/0.18\ \mu\text{m}$$

The minimum required width of the set pulse is composed of two components: the time for $v_{\overline{Q}}$ in the circuit of Fig. 16.5(a) to fall from V_{DD} to $V_{DD}/2$, where $V_{DD}/2$ is the threshold voltage of the inverter formed by Q_3 and Q_4 in Fig. 16.4, and the time for the output of the Q_3–Q_4 inverter to rise from 0 to $V_{DD}/2$. At the end of the second time interval, the feedback signal will have traveled around the feedback loop, and regeneration can continue without the presence of the set pulse. We will denote the first component t_{PHL} and the second t_{PLH}, and will calculate their values as follows.

To determine t_{PHL} refer to the circuit in Fig. 16.6 and note that the capacitor discharge current i_C is the difference between the current of the equivalent transistor Q_{eq} and the current of Q_2,

$$i_C = i_{D\text{eq}} - i_{D2}$$

To determine the average discharge current i_C, we calculate $i_{D\text{eq}}$ and i_{D2} at $t = 0$ and $t = t_{PHL}$. At $t = 0$, $v_{\overline{Q}} = V_{DD}$, thus Q_2 is off,

$$i_{D2}(0) = 0$$

and Q_{eq} is in saturation,

$$i_{D\text{eq}} = \frac{1}{2} \times 300 \times \frac{1}{2} \times \frac{1.08}{0.18} \times (1.8 - 0.5)^2$$
$$= 760.5\ \mu\text{A}$$

Thus,

$$i_C(0) = 760.5 - 0 = 760.5\ \mu\text{A}$$

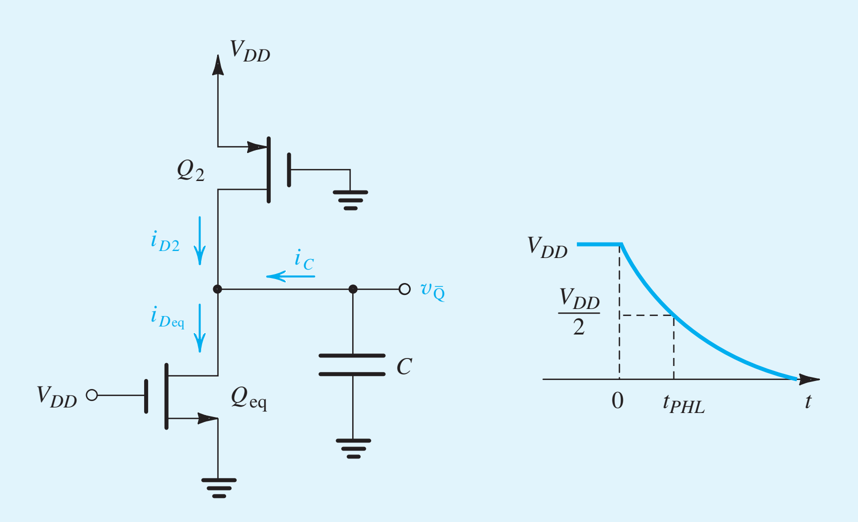


Figure 16.6 Determining the time t_{PHL} for $v_{\bar{Q}}$ to fall from V_{DD} to $V_{DD}/2$.

At $t = t_{PHL}$, $v_{\bar{Q}} = V_{DD}/2$, thus both Q_2 and Q_{eq} will be in the triode region,

$$i_{D2}(t_{PHL}) = 75 \times \frac{1.08}{0.18} \times \left[(1.8 - 0.5)\left(\frac{1.8}{2}\right) - 0.5\left(\frac{1.8}{2}\right)^2\right]$$
$$= 344.25\ \mu\text{A}$$

and

$$i_{Deq}(t_{PHL}) = 300 \times \frac{1}{2} \times \frac{1.08}{0.18}\left[(1.8 - 0.5)\left(\frac{1.8}{2}\right) - 0.5\left(\frac{1.8}{2}\right)^2\right]$$
$$= 688.5\ \mu\text{A}$$

Thus,

$$i_C(t_{PHL}) = 688.5 - 344.25 = 344.25\ \mu\text{A}$$

and the average value of i_C over the interval $t = 0$ to $t = t_{PHL}$ is

$$i_C\big|_{\text{av}} = \frac{i_C(0) + i_C(t_{PHL})}{2}$$
$$= \frac{760.5 + 344.25}{2} = 552.4\ \mu\text{A}$$

Example 16.1 *continued*

We now can calculate t_{PHL} as

$$t_{PHL} = \frac{C(V_{DD}/2)}{i_C|_{\text{av}}} = \frac{20 \times 10^{-15} \times 0.9}{552.4 \times 10^{-6}} = 32.6 \text{ ps}$$

Next we consider the time t_{PHL} for the output of the Q_3-Q_4 inverter, v_Q, to rise from 0 to $V_{DD}/2$. The value of t_{PLH} can be calculated using the propagation-delay formula derived in Chapter 14 (Eq. 14.52), which is also listed in Table 14.2, namely,

$$t_{PLH} = \frac{\alpha_p C}{k_p'(W/L)_p V_{DD}}$$

where

$$\alpha_p = 2 \Bigg/ \left[\frac{7}{4} - \frac{3|V_{tp}|}{V_{DD}} + \left(\frac{|V_{tp}|}{V_{DD}}\right)^2\right]$$

Substituting numerical values we obtain,

$$\alpha_p = \frac{2}{1.75 - \dfrac{3 \times 0.5}{1.8} + \left(\dfrac{0.5}{1.8}\right)^2} = 2.01$$

and

$$t_{PLH} = \frac{2.01 \times 20 \times 10^{-15}}{75 \times 10^{-6} \times (1.08/0.18) \times 1.8} = 49.7 \text{ ps}$$

Finally, the minimum required width of the set pulse can be calculated as

$$T_{\min} = t_{PHL} + t_{PLH} = 82.3 \text{ ps}$$

EXERCISE

16.1 For the SR flip-flop specified in Example 16.1, find the minimum W/L for both Q_5 and Q_6 so that switching is achieved when inputs S and ϕ are at $(V_{DD}/2)$.
Ans. 28.7

16.1.4 A Simpler CMOS Implementation of the Clocked SR Flip-Flop

A simpler implementation of a clocked SR flip-flop is shown in Fig. 16.7. Here, pass-transistor logic is employed to implement the clocked set–reset functions. This circuit is very popular in the design of static random-access memory (SRAM) chips, where it is used as the basic memory cell (Section 16.4.1).

16.1.5 D Flip-Flop Circuits

A variety of flip-flop types exist and can be synthesized using logic gates. CMOS circuit implementations can be obtained by simply replacing the gates with their CMOS circuit realizations. This approach, however, usually results in rather complex circuits. In many cases, simpler circuits can be found by taking a circuit-design viewpoint, rather than a logic-design one. To illustrate this point, we shall consider the CMOS implementation of a very important type of flip-flop, the data, or D, flip-flop.

The D flip-flop is shown in block diagram form in Fig. 16.8. It has two inputs, the data input D and the clock input ϕ. The complementary outputs are labeled Q and $\overline{Q}$. When the clock is low, the flip-flop is in the memory, or rest, state; signal changes on the D input line have no effect on the state of the flip-flop. As the clock goes high, the flip-flop acquires the logic level that existed on the D line just before the rising edge of the clock. Such a flip-flop is said to be **edge triggered**. Some implementations of the D flip-flop include direct set and reset inputs that override the clocked operation just described.

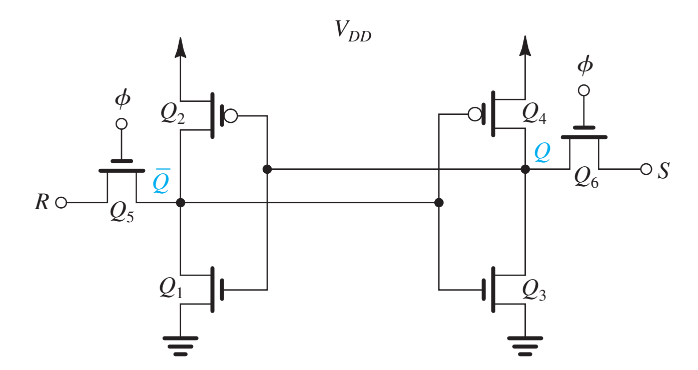


Figure 16.7 A simpler CMOS implementation of the clocked SR flip-flop. This circuit is popular as the basic cell in the design of static random-access memory (SRAM) chips.

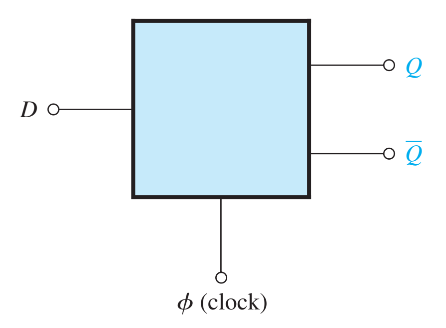


Figure 16.8 A block diagram representation of the D flip-flop.

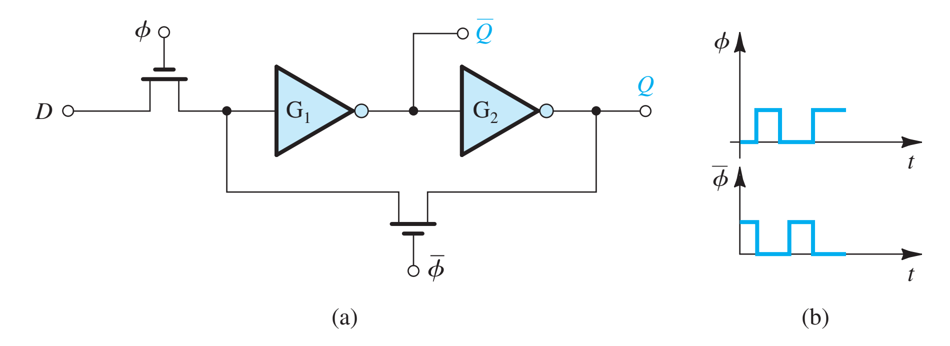


Figure 16.9 A simple implementation of the D flip-flop. The circuit in **(a)** utilizes the two-phase **nonoverlapping clock** whose waveforms are shown in **(b).**

A simple implementation of the D flip-flop is shown in Fig. 16.9. The circuit consists of two inverters connected in a positive-feedback loop, just as in the static latch of Fig. 16.1(a), except that here the loop is closed for only part of the time. Specifically, the loop is closed when the clock is low ($\phi = 0, \overline{\phi} = 1$). The input D is connected to the flip-flop through a switch that closes when the clock is high. Operation is straightforward: When ϕ is high, the loop is opened, and the input D is connected to the input of inverter G_1. The capacitance at the input node of G_1 is charged to the value of D, and the capacitance at the input node of G_2 is charged to the value of $\overline{D}$. Then, when the clock goes low, the input line is isolated from the flip-flop, the feedback loop is closed, and the latch acquires the state corresponding to the value of D just before ϕ went down, providing an output $Q = D$.

From the preceding, we observe that the circuit in Fig. 16.9 combines the positive-feedback technique of static bistable circuits and the charge-storage technique of dynamic circuits. It is important to note that the proper operation of this circuit, and of many circuits that use clocks, is predicated on the assumption that *ϕ and $\overline{\phi}$ will not be simultaneously high at any time*. This condition is defined by referring to the two clock phases as being *nonoverlapping*.

An inherent drawback of the D flip-flop implementation of Fig. 16.9 is that during ϕ, the output of the flip-flop simply follows the signal on the D input line. This can cause problems in certain logic-design situations. The problem is solved very effectively by using the **master–slave** configuration shown in Fig. 16.10(a). Before discussing its circuit operation, we note that although the switches are shown implemented with single NMOS transistors, CMOS transmission gates are employed in many applications. We are simply using the single MOS transistor as a "shorthand notation" for a series switch.

The master–slave circuit consists of a pair of circuits of the type shown in Fig. 16.9, operated with alternate clock phases. Here, to emphasize that the two clock phases must be nonoverlapping, we denote them ϕ_1 and ϕ_2, and clearly show the nonoverlap interval in the waveforms of Fig. 16.10(b). Operation of the circuit is as follows:

1. When ϕ_1 is high and ϕ_2 is low, the input is connected to the master latch, whose feedback loop is opened, while the slave latch is isolated. Thus, the output Q remains at the value stored previously in the slave latch, whose loop is now closed. The node capacitances of the master latch are charged to the appropriate voltages corresponding to the present value of D.
2. When ϕ_1 goes low, the master latch is isolated from the input data line. Then, when ϕ_2 goes high, the feedback loop of the master latch is closed, locking in the value of D. Further, its output is connected to the slave latch, whose feedback loop is now

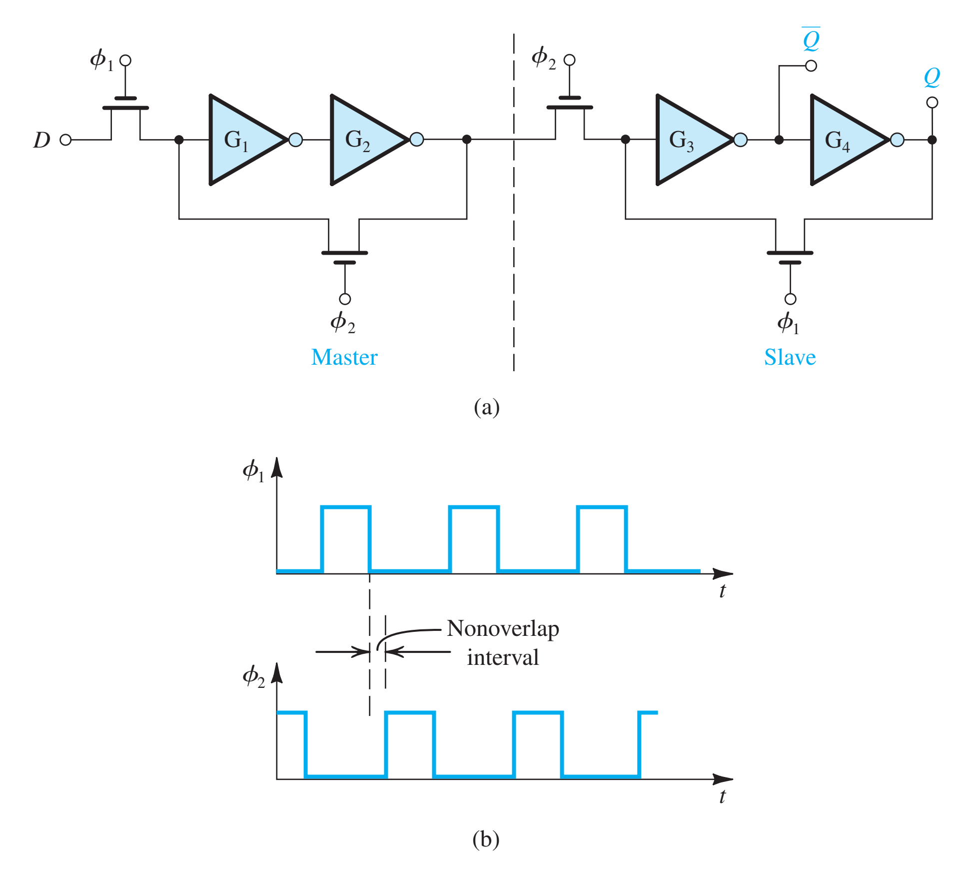


Figure 16.10 **(a)** A master–slave D flip-flop. The switches can be, and usually are, implemented with CMOS transmission gates. **(b)** Waveforms of the two-phase nonoverlapping clock required.

open. The node capacitances in the slave are appropriately charged so that when ϕ_1 goes high again, the slave latch locks in the new value of D and provides it at the output, $Q = D$.

From this description, we note that at the positive transition of clock ϕ_2 the output Q adopts the value of D that existed on the D line at the end of the preceding clock phase, ϕ_1. This output value remains constant for one clock period. Finally, note that during the nonoverlap interval both latches have their feedback loops open, and we are relying on the node capacitances to maintain most of their charge. It follows that the nonoverlap interval should be kept reasonably short (perhaps one-tenth or less of the clock period, and of the order of 1 ns or so in current practice).

16.2 Semiconductor Memories: Types and Architectures

A computer system, whether a large machine or a microcomputer, requires memory for storing data and program instructions. Furthermore, within a given computer system there usually are various types of memory utilizing a variety of technologies and having different *access*

times. Broadly speaking, computer memory can be divided into two types: **main memory** and **mass-storage** memory. The main memory is usually the most rapidly accessible memory and the one from which most, often all, instructions in programs are executed. The main memory is usually of the random-access type. A **random-access memory** (RAM) is one in which the time required for storing (writing) information and for retrieving (reading) information is independent of the physical location (within the memory) in which the information is stored.

Random-access memories should be contrasted with *serial* or *sequential* memories, such as disks and tapes, from which data are available only in the sequence in which the data were originally stored. Thus, in a serial memory the time to access particular information depends on the memory location in which the required information is stored, and the average access time is longer than the access time of random-access memory. In a computer system, serial memory is used for mass storage. Items not frequently accessed, such as large parts of the computer operating system, are usually stored in a *moving-surface memory* such as magnetic disk.

Another important classification of memory relates to whether it is a **read/write** or a **read-only memory**. Read/write (R/W) memory permits data to be stored and retrieved at comparable speeds. Computer systems require random-access read/write memory for data and program storage.

Read-only memories (**ROM**) permit reading at the same high speeds as R/W memories (or perhaps higher) but restrict the writing operation. ROMs can be used to store a microprocessor operating-system program. They are also employed in operations that require table lookup, such as finding the values of mathematical functions. A popular application of ROMs is their use in video game cartridges. It should be noted that read-only memory is usually of the random-access type. Nevertheless, in the digital circuit jargon, the acronym RAM usually refers to read/write, random-access memory, while ROM is used for read-only memory.

The regular structure of memory circuits has made them an ideal application for the design of circuits of the very-large-scale integrated (VLSI) type. Indeed, at any moment, memory chips represent the state of the art in packing density and hence integration level. Beginning with the introduction of the 1-Kbit chip in 1970, memory-chip density has quadrupled about every 3 years. At the present time (2013), chips containing 4 Gbit[1] are available. In this and the next two sections, we shall study some of the basic circuits employed in VLSI RAM chips. Read-only memory circuits are studied in Section 16.5.

16.2.1 Memory-Chip Organization

The bits on a memory chip are addressable either individually or in groups of 4 to 16. As an example, a 64-Mbit chip in which all bits are individually addressable is said to be organized as 64M words × 1 bit (or simply 64M × 1). Such a chip needs a 26-bit address ($2^{26} = 67{,}108{,}864 = 64\text{M}$). On the other hand, the 64-Mbit chip can be organized as 16M words × 4 bits (16M × 4), in which case a 24-bit address is required. For simplicity we shall assume in our subsequent discussion that all the bits on a memory chip are individually addressable.

The bulk of the memory chip consists of the cells in which the bits are stored. Each **memory cell** is an electronic circuit capable of storing one bit. We shall study memory-cell circuits in Section 16.3. For reasons that will become clear shortly, it is desirable to physically organize

[1] The capacity of a memory chip to hold binary information as binary digits (or bits) is measured in kilobit (Kbit), megabit (Mbit), and gigabit (Gbit) units, where 1 Kbit = 1024 bits, 1 Mbit = 1024 × 1024 = 1,048,576 bits, and 1 Gbit = 1024^3 bits. Thus a 64-Mbit chip contains 67,108,864 bits of memory.

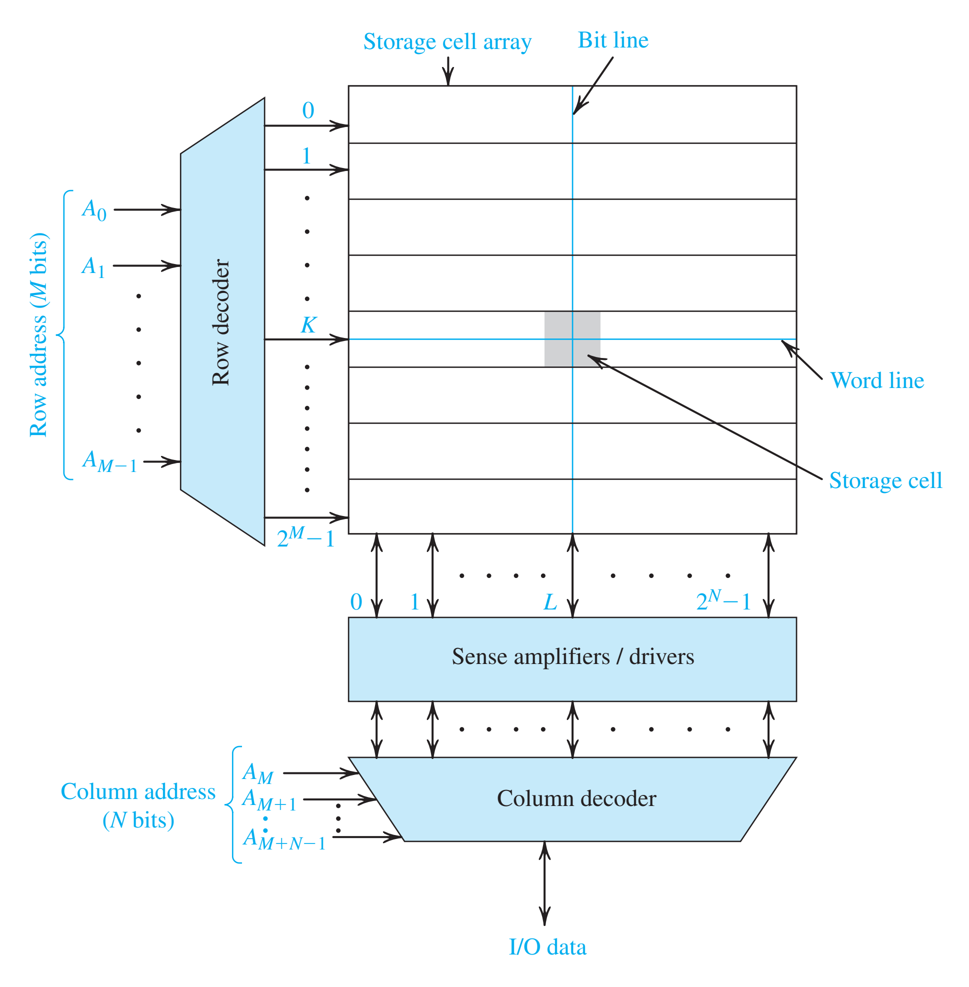


Figure 16.11 A 2^{M+N}-bit memory chip organized as an array of 2^M rows $\times$ 2^N columns.

the storage cells on a chip in a square or a nearly square matrix. Figure 16.11 illustrates such an organization. The cell matrix has 2^M rows and 2^N columns, for a total storage capacity of 2^{M+N}. For example, a 1M-bit square matrix would have 1024 rows and 1024 columns ($M = N = 10$). Each cell in the array is connected to one of the 2^M row lines, known rather loosely, but universally, as **word lines**, and to one of the 2^N column lines, known as **digit lines** or, more commonly, **bit lines**. A particular cell is **selected** for reading or writing by activating its word line and its bit line.

Activating one of the 2^M word lines is performed by the **row decoder**, a combinational logic circuit that selects (raises the voltage of) the particular word line whose M-bit address is applied to the decoder input. The address bits are denoted $A_0, A_1, \ldots, A_{M-1}$. When the Kth word line is activated for, say, **a read operation**, all 2^N cells in row K will provide their contents to their respective bit lines. Thus, if the cell in column L (Fig. 16.11) is storing a 1, the voltage of bit-line number L will be raised, usually by a small voltage, say 0.1 V to 0.2 V. The readout voltage is small because the cell is small, a deliberate design decision, since the number of cells is very large. The small readout signal is applied to a **sense amplifier** connected to the bit line. As Fig. 16.11 indicates, there is a sense amplifier for every bit line. The sense amplifier

provides a full-swing digital signal (from 0 to V_{DD}) at its output. This signal, together with the output signals from all the other cells in the selected row, is then delivered to the **column decoder**. The column decoder selects the signal of the particular column whose N-bit address is applied to the decoder input (the address bits are denoted $A_M, A_{M+1}, \ldots, A_{M+N-1}$) and causes this signal to appear on the chip input/output (I/O) data line.

A **write operation** proceeds in a similar manner: The data bit to be stored (1 or 0) is applied to the I/O line. The cell in which the data bit is to be stored is selected through the combination of its row address and its column address. The sense amplifier of the selected column acts as a **driver** to write the applied signal into the selected cell. Circuits for sense amplifiers and address decoders will be studied in Section 16.4.

Before leaving the topic of memory organization (or memory-chip architecture), we wish to mention a relatively recent innovation in organization dictated by the exponential increase in chip density. To appreciate the need for a change, note that as the number of cells in the array increases, the physical lengths of the word lines and the bit lines increase. This has occurred even though for each new generation of memory chips, the transistor size has decreased (currently, CMOS process technologies with 22-nm feature size are utilized). The net increase in word-line and bit-line lengths increases their total resistance and capacitance, and thus slows down their transient response. That is, as the lines lengthen, the exponential rise of the voltage of the word line becomes slower, and it takes longer for the cells to be activated. This problem has been solved by partitioning the memory chip into a number of blocks. Each of the blocks has an organization identical to that in Fig. 16.11. The row and column addresses are broadcast to all blocks, but the data selected come from only one of the blocks. Block selection is achieved by using an appropriate number of the address bits as a block address. Such an architecture can be thought of as three-dimensional: rows, columns, and blocks.

16.2.2 Memory-Chip Timing

The **memory access time** is the time between the initiation of a read operation and the appearance of the output data. The **memory cycle time** is the minimum time allowed between two consecutive memory operations. To be on the conservative side, a memory operation is usually taken to include both read and write (in the same location). MOS memories have access and cycle times in the range of a few to a few hundred nanoseconds.

EXERCISES

16.2 A 4-Mbit memory chip is partitioned into 32 blocks, with each block having 1024 rows and 128 columns. Give the number of bits required for the row address, column address, and block address.
Ans. 10; 7; 5

16.3 The word lines in a particular MOS memory chip are fabricated using polysilicon (see Appendix A). The resistance of each word line is estimated to be 5 kΩ, and the total capacitance between the line and ground is 2 pF. Find the time for the voltage on the word line to reach $V_{DD}/2$, assuming that the line is driven by a voltage V_{DD} provided by a low-impedance inverter. (*Note:* The line is actually a distributed network that we are approximating by means of a lumped circuit consisting of a single resistor and a single capacitor.)
Ans. 6.9 ns

16.3 Random-Access Memory (RAM) Cells

As mentioned in Section 16.2, the major part of the memory chip is taken up by the storage cells. It follows that to be able to pack a large number of bits on a chip, it is imperative that the cell be reduced to the smallest size possible. The power dissipation per cell should be minimized also. Thus, many of the flip-flop circuits studied in Section 16.1 are too complex to be suitable for implementing the storage cells in a RAM chip.

There are basically two types of MOS RAM: static and dynamic. **Static RAMs** (called **SRAMs** for short) utilize static latches as the storage cells. Dynamic RAMs (called **DRAMs**), on the other hand, store the binary data on capacitors, resulting in further reduction in cell area, but at the expense of more complex read and write circuitry. In particular, while static RAMs can hold their stored data indefinitely, provided the power supply remains on, dynamic RAMs require *periodic refreshing* to regenerate the data stored on capacitors. This is because the storage capacitors will discharge, though slowly, as a result of the leakage currents inevitably present. By virtue of their smaller cell size, dynamic memory chips are usually four times as dense as their contemporary static chips. Thus while the state of the art in 2013 is a 4-Gbit DRAM chip, the highest-density SRAM chip has 1-Gbit capacity. Both static and dynamic RAMs are *volatile*; that is, they require the continuous presence of a power supply. By contrast, most ROMs are of the nonvolatile type, as we shall see in Section 16.5. In the following subsections, we shall study basic SRAM and DRAM storage cells.

16.3.1 Static Memory (SRAM) Cell

Figure 16.12 shows a typical static memory cell in CMOS technology. The circuit, which we encountered in Section 16.1, is a flip-flop comprising two cross-coupled inverters and two **access transistors**, Q_5 and Q_6. The access transistors are turned on when the word line is selected and its voltage raised to V_{DD}, and they connect the flip-flop to the column (bit or B) line and $\overline{\text{column}}$ ($\overline{\text{bit}}$ or $\overline{B}$) line. Note that although in principle only the B or the $\overline{B}$ line suffices, most often both are utilized, as shown in Fig. 16.12. This practice both provides a

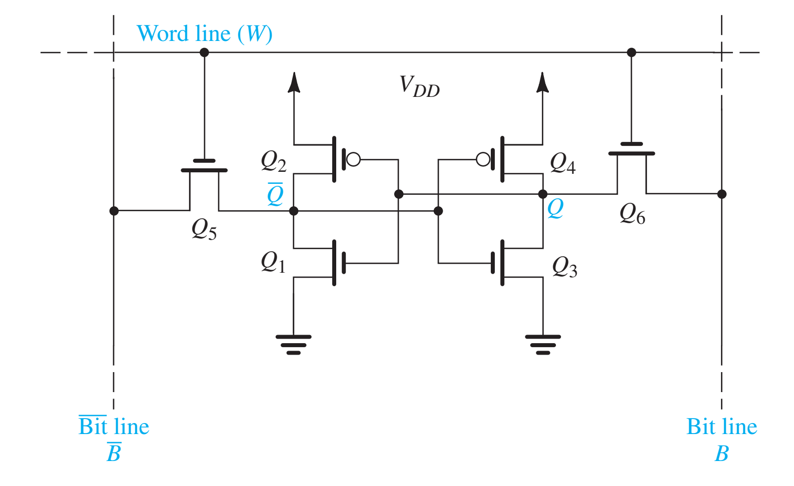


Figure 16.12 A CMOS SRAM memory cell.

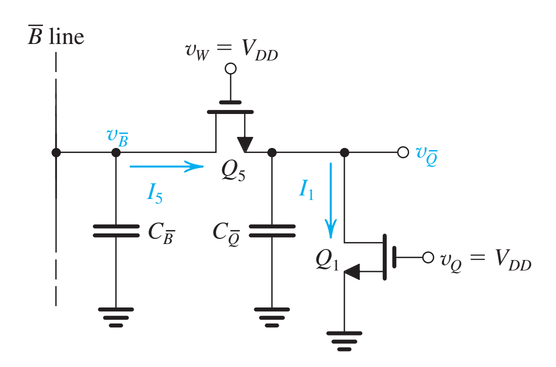


Figure 16.13 Relevant parts of the SRAM cell circuit during a read operation when the cell is storing a logic 1. Note that initially $v_Q = V_{DD}$ and $v_{\overline{Q}} = 0$. Also note that the B and $\overline{B}$ lines are precharged to a voltage V_{DD}.

differential data path between the cell and the memory-chip output and increases the circuit reliability. The access transistors act as transmission gates allowing bidirectional current flow between the flip-flop and the B and $\overline{B}$ lines. Finally, we note that this circuit is known as the **six-transistor** or **6T cell**.

The Read Operation Consider first a read operation, and assume that the cell is storing a 1. In this case, Q will be high at V_{DD}, and $\overline{Q}$ will be low at 0 V. Before the read operation begins, the B and $\overline{B}$ lines are raised to a voltage in the range $V_{DD}/2$ to V_{DD}. This process, known as **precharging**, is performed using circuits we shall discuss in Section 16.4 in conjunction with the study of sense amplifiers. To simplify matters, we shall assume here that the precharge voltage of B and $\overline{B}$ is V_{DD}.

When the word line is selected and the access transistors Q_5 and Q_6 are turned on, examination of the circuit reveals that the only portion that will be conducting is that shown in Fig. 16.13. Noting that the initial value of $v_{\overline{Q}}$ is 0 V, we can see that current will flow from the $\overline{B}$ line (actually, from the $\overline{B}$-line capacitance $C_{\overline{B}}$) through Q_5 and into capacitor $C_{\overline{Q}}$, which is the small equivalent capacitance between the $\overline{Q}$ node and ground. This current charges $C_{\overline{Q}}$ and thus $v_{\overline{Q}}$ rises and Q_1 conducts, sinking some of the current supplied by Q_5. Equilibrium will be reached when $C_{\overline{Q}}$ is charged to a voltage $V_{\overline{Q}}$ at which I_1 equals I_5, and no current flows through $C_{\overline{Q}}$. Here it is extremely important to note that to avoid changing the state of the flip-flop, that is, for our read operation to be **nondestructive**, $V_{\overline{Q}}$ must not exceed the threshold voltage of the inverter Q_3–Q_4. In fact, SRAM designers usually impose a more stringent requirement on the value of $V_{\overline{Q}}$, namely, that it should be lower than the threshold voltage of Q_3, V_{tn}. Thus, the design problem we shall now solve is as follows: Determine the ratio of $(W/L)_5/(W/L)_1$ so that $V_{\overline{Q}} \leq V_{tn}$.

Noting that Q_5 will be operating in saturation and neglecting, for simplicity, the body effect, we can write

$$I_5 = \frac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)_5 (V_{DD} - V_{tn} - V_{\overline{Q}})^2 \tag{16.1}$$

Transistor Q_1 will be operating in the triode region, and its current I_1 can be written as

$$I_1 = (\mu_n C_{ox})\left(\frac{W}{L}\right)_1 \left[(V_{DD} - V_{tn})V_{\overline{Q}} - \frac{1}{2}V_{\overline{Q}}^2\right] \tag{16.2}$$

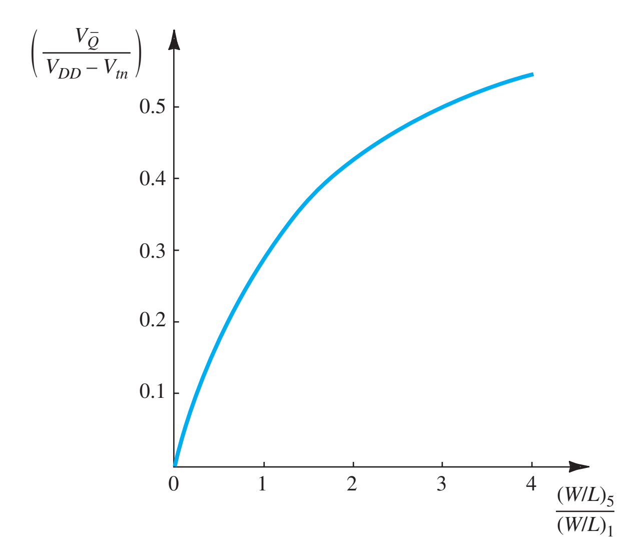


Figure 16.14 The normalized value of $V_{\bar{Q}}$ versus the ratio $(W/L)_5/(W/L)_1$ for the circuit in Fig. 16.13. This graph can be used to determine the maximum value permitted for $(W/L)_5/(W/L)_1$ so that $V_{\bar{Q}}$ is kept below a desired level.

Equating I_5 and I_1 gives a quadratic equation in $V_{\bar{Q}}$, which can be solved to obtain

$$\frac{V_{\bar{Q}}}{V_{DD}-V_{tn}} = 1 - 1 \Bigg/ \sqrt{1 + \frac{(W/L)_5}{(W/L)_1}} \qquad (16.3)$$

This is an attractive relationship, since it provides $V_{\bar{Q}}$ in normalized form and thus always applies, independent of the process technology utilized. Figure 16.14 shows a universal plot of $\left[V_{\bar{Q}}/(V_{DD}-V_{tn})\right]$ versus $(W/L)_5/(W/L)_1$. For a given process technology, V_{DD} and V_{tn} are determined, and the plot in Fig. 16.14 can be used to determine the maximum value permitted for $(W/L)_5/(W/L)_1$ while keeping $V_{\bar{Q}}$ below a desired value. Alternatively, we can derive a formula for this purpose. For instance, if $V_{\bar{Q}}$ is to be kept below V_{tn}, the ratio of $(W/L)_5$ to $(W/L)_1$ must be kept below the value obtained from Eq. (16.3), that is,

$$\frac{(W/L)_5}{(W/L)_1} \leq \frac{1}{\left(1 - \dfrac{V_{tn}}{V_{DD}-V_{tn}}\right)^2} - 1 \qquad (16.4)$$

This is an important design constraint that can be expressed in a slightly more general form by replacing $(W/L)_5$ with $(W/L)_a$, where the subscript a denotes access transistors Q_5 and Q_6, and $(W/L)_1$ with $(W/L)_n$, which is the W/L ratio of Q_N in each of the two inverters; thus,

$$\frac{(W/L)_a}{(W/L)_n} \leq \frac{1}{\left(1 - \dfrac{V_{tn}}{V_{DD}-V_{tn}}\right)^2} - 1 \qquad (16.5)$$

EXERCISE

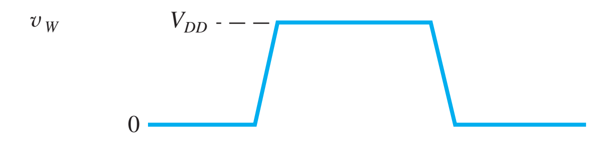
16.4 Find the maximum allowable W/L for the access transistors of the SRAM cell in Fig. 16.12 so that in a read operation, the voltages at Q and $\overline{Q}$ do not change by more than $|V_t|$. Assume that the SRAM is fabricated in a 0.18-μm technology for which $V_{DD} = 1.8$ V, $V_{tn} = |V_{tp}| = 0.5$ V and that $(W/L)_n = 1.5$.
Ans. $(W/L)_a \leq 2.5$

Having determined the constraint imposed by the read operation on the W/L ratios of the access transistors, we now return to the circuit in Fig. 16.13, and show in Fig. 16.15 the voltage waveforms at various nodes during a read-1 operation. Observe that as we have already discussed, $v_{\overline{Q}}$ rises from zero to a voltage $V_{\overline{Q}} \leq V_{tn}$. Correspondingly, the change in v_Q will be very small, justifying the assumption implicit in the analysis above that v_Q remains constant at V_{DD}. Most important, note that the voltage of the $\overline{B}$ line, $v_{\overline{B}}$, decreases by a small amount ΔV. This is a result of the discharge of the capacitance of the $\overline{B}$ line, $C_{\overline{B}}$, by the current I_5. Assuming that I_5 reaches its equilibrium value in Eq. (16.1) relatively quickly, capacitor $C_{\overline{B}}$ is in effect discharged by a constant current I_5 and the change in its voltage, ΔV, obtained in a time interval Δt, can be found by writing a charge-balance equation,

$$I_5 \Delta t = C_{\overline{B}} \Delta V$$

Thus,

$$\Delta V = \frac{I_5 \Delta t}{C_{\overline{B}}} \tag{16.6}$$









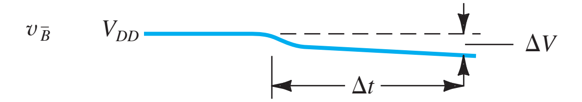


Figure 16.15 Voltage waveforms at various nodes in the SRAM cell during a read-1 operation.

Here we note that $C_{\overline{B}}$ is usually relatively large (1–2 pF) because a large number of cells are connected to the $\overline{B}$ line. The incremental change ΔV is therefore rather small (0.1–0.2 V), necessitating the use of a sense amplifier. If the sense amplifier requires a minimum decrement ΔV in $v_{\overline{B}}$ to detect the presence of a "1," then the read delay time can be found from Eq. (16.6) as

$$\Delta t = \frac{C_{\overline{B}}\Delta V}{I_5} \tag{16.7}$$

This equation indicates the need for a relatively large I_5 to reduce the delay time Δt. A large I_5, however, implies selecting $(W/L)_a$ near the upper bound given by Eq. (16.5), which in turn means an increase in the silicon area occupied by the access transistors and hence the cell area, an interesting design trade-off.

EXERCISE

16.5 For the SRAM cell considered in Exercise 16.4 whose $(W/L)_n = 1.5$ and $(W/L)_a \leq 2.5$, use Eq. (16.7) to determine the read delay Δt in two cases: (a) $(W/L)_a = 2.5$ and (b) $(W/L)_a = 1.5$. Let $\mu_n C_{ox} = 300\ \mu\text{A/V}^2$. In both cases, assume that $C_{\overline{B}} = 2$ pF and that the sense amplifier requires a ΔV of minimum magnitude of 0.2 V. (*Hint*: Use Eq. 16.1 to determine I_5, and Eq. 16.3 to determine $V_{\overline{Q}}$.)
Ans. 1.7 ns; 2.1 ns

We conclude our discussion of the read operation with two remarks:

1. Although we considered only the read-1 operation, the read-0 operation is identical; it involves Q_3 and Q_6 with the analysis resulting in an upper bound on $(W/L)_6/(W/L)_3$ equal to that we have found for $(W/L)_5/(W/L)_1$. This, of course, is entirely expected, since the circuit is symmetrical. The read-0 operation results in a decrement ΔV in the voltage of the B line, which is interpreted by the sense amplifier as a stored 0.
2. The component Δt of the read delay is relatively large because C_B and $C_{\overline{B}}$ are relatively large (in the picofarad range). Also, Δt is not the only component of the read delay; another significant component is due to the finite rise time of the voltage on the word line. Indeed, even the calculation of Δt is optimistic, since the word line will have only reached a voltage lower than V_{DD} when the process of discharging $C_{\overline{B}}$ takes place. As will be seen shortly, the write operation is faster.

The Write Operation We next consider the write operation. Let the SRAM cell of Fig. 16.12 be storing a logic 1, thus $V_Q = V_{DD}$ and $V_{\overline{Q}} = 0$ V, and assume that we wish to write a 0; that is, we wish to have the flip-flop switch states. To write a zero, the B line is lowered to 0 V, and the $\overline{B}$ line is raised to V_{DD} and, of course, the cell is selected by raising the word line to V_{DD}. The objective now is to pull node Q down and node $\overline{Q}$ up and have the voltage of at least one of these two nodes pass by the inverter threshold voltage. Thus, if v_Q decreases below the threshold voltage of inverter Q_1–Q_2, the regenerative action of the latch will start and the flip-flop will switch to the stored-0 state. Alternatively, or in addition, if we manage to raise $v_{\overline{Q}}$ above the threshold voltage of the Q_3–Q_4 inverter, the regenerative action

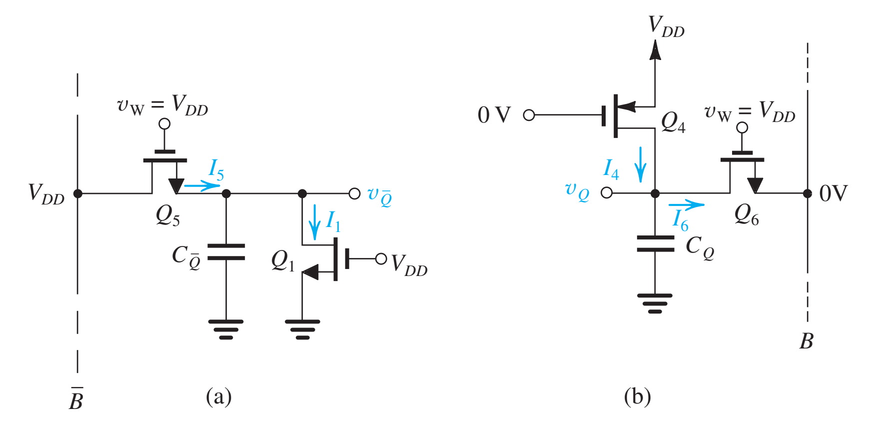


Figure 16.16 Relevant parts of the 6T SRAM circuit of Fig. 16.12 during the process of writing a 0. It is assumed that the cell is originally storing a 1 and thus initially $v_Q = V_{DD}$ and $v_{\bar{Q}} = 0$ V.

will be engaged and the latch will eventually switch state. Either one of the two actions is sufficient to engage the regenerative mechanism of the latch.

Figure 16.16 shows the relevant parts of the SRAM circuit during the interval when $v_{\bar{Q}}$ is being pulled up [Fig. 16.16(a)] and v_Q is being pulled down [Fig. 16.16(b)]. Since **toggling** (i.e., state change) has not yet taken place, we assume that the voltage feeding the gate of Q_1 is still equal to V_{DD} and the voltage at the gate of Q_4 is still equal to 0 V. These voltages will of course be changing as $v_{\bar{Q}}$ goes up and v_Q goes down, but this assumption is nevertheless reasonable for approximate hand analysis.

Consider first the circuit in Fig. 16.16(a). This is the same circuit we analyzed in detail in the study of the read operation above. Recall that to make the read process nondestructive, we imposed an upper bound on $(W/L)_5$. That upper bound ensured that $v_{\bar{Q}}$ will not rise above V_{tn}. Thus, this circuit is not capable of raising $v_{\bar{Q}}$ to the point that it can start the regenerative action. We must therefore rely solely on the circuit of Fig. 16.16(b). That is, our write-0 operation will be accomplished by pulling node Q down in order to initiate the regenerative action of the latch. To ensure that the latch will in fact switch state, SRAM designers impose a more stringent requirement on the voltage v_Q, namely, that it must fall below not just V_M of the Q_1–Q_2 inverter but below V_{tn} of Q_1.

Let's now look more closely at the circuit of Fig. 16.16(b). Initially, v_Q is at V_{DD}. However, as Q_6 turns on, I_6 quickly discharges the small capacitance C_Q, and v_Q begins to fall. This will enable Q_4 to conduct, and equilibrium is reached when $I_4 = I_6$. To ensure toggling, we design the circuit so that this equilibrium occurs at a value of v_Q less than V_{tn}. At such a value V_Q, Q_4 will be operating in saturation (or at least at the edge of saturation) and Q_6 will be operating in the triode region, thus

$$I_4 = \frac{1}{2}(\mu_p C_{ox})\left(\frac{W}{L}\right)_4 (V_{DD} - |V_{tp}|)^2 \tag{16.8}$$

and

$$I_6 = (\mu_n C_{ox})\left(\frac{W}{L}\right)_6 \left[(V_{DD} - V_{tn})V_Q - \frac{1}{2}V_Q^2\right] \tag{16.9}$$

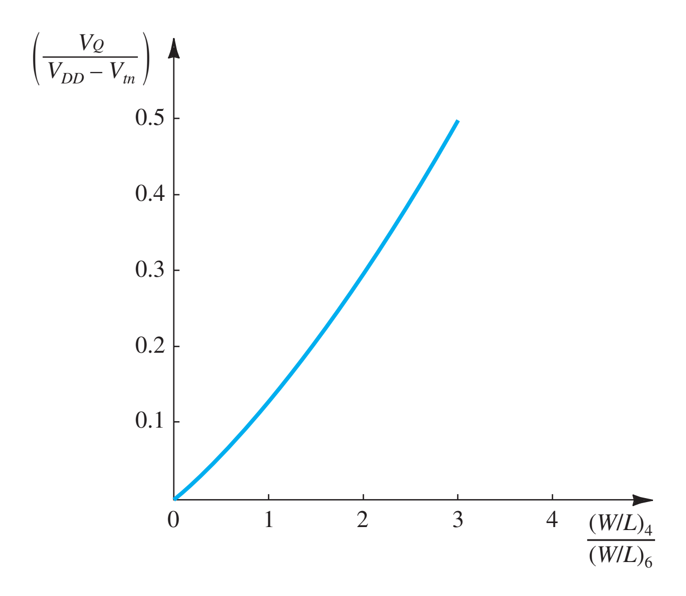


Figure 16.17 The normalized value of V_Q versus the ratio $(W/L)_4/(W/L)_6$ for the circuit in Fig. 16.16(b). The graph applies for process technologies for which $\mu_n \simeq 4\mu_p$. It can be used to determine the maximum $(W/L)_4/(W/L)_6$ for which V_Q is guaranteed to fall below a desired value.

Substituting $|V_{tp}| = V_{tn}$, which is usually the case, and equating I_4 and I_6 results in a quadratic equation in V_Q whose solution is

$$\frac{V_Q}{V_{DD} - V_{tn}} = 1 - \sqrt{1 - \left(\frac{\mu_p}{\mu_n}\right)\frac{(W/L)_4}{(W/L)_6}} \tag{16.10}$$

This relationship is not as convenient as that in Eq. (16.3) because the right-hand side includes a process-dependent quantity, namely, μ_p/μ_n. Thus we do not have a universally applicable relationship. Nevertheless, for a number of CMOS process technologies, including the 0.25-μm, the 0.18-μm, and the 0.13-μm processes, $\mu_n/\mu_p \simeq 4$. Thus, upon substituting $\mu_p/\mu_n = 0.25$ in Eq. (16.10), we obtain the semiuniversal graph shown in Fig. 16.17. We can use this graph to determine the maximum allowable value of the ratio $(W/L)_4/(W/L)_6$ that will ensure a value of $V_Q \leq V_{tn}$ for given process parameters V_{DD} and V_{tn}. Alternatively, substituting $V_Q = V_{tn}$, $(W/L)_4 = (W/L)_p$, and $(W/L)_6 = (W/L)_a$, we can obtain the upper bound analytically as

$$\frac{(W/L)_p}{(W/L)_a} \leq \left(\frac{\mu_n}{\mu_p}\right)\left[1 - \left(1 - \frac{V_{tn}}{V_{DD} - V_{tn}}\right)^2\right] \tag{16.11}$$

Observe that this relationship provides an upper bound on $(W/L)_p$ in terms of $(W/L)_a$ and that the relationship in Eq. (16.5) provides an upper bound on $(W/L)_a$ in terms of $(W/L)_n$. Thus, the two relationships can be used together to design the SRAM cell.

EXERCISE

D16.6 For the SRAM cell considered in Exercise 16.4, where $(W/L)_n = 1.5$ and $(W/L)_a \leq 2.5$, use Eq. (16.11) to find the maximum allowable value of $(W/L)_p$. Recall that for this 0.18-μm process, $\mu_n \simeq 4\mu_p$. For all transistors having $L = 0.18$ μm, find W_n, W_p, and W_a that result in a minimum-area cell. Assume that the minimum allowable width is 0.18 μm.
Ans. $(W/L)_p \leq 2.5(W/L)_a$, thus $(W/L)_p \leq 6.25$; for minimum area, select $W_n = W_p = W_a = 0.18$ μm.

We conclude our study of the write process by noting that it is fast because it does not require discharging the large capacitance of the bit lines. The voltages of the B and $\overline{B}$ lines are driven to their required values of 0 or V_{DD} by powerful driver circuits and thus achieve their desired voltages very quickly. The write delay is determined roughly by the time for the regenerating signal to propagate around the feedback loop of the latch; thus it is about twice the propagation delay of the inverter. Of course, the write cycle time is still lengthened by the word-line delay.

16.3.2 Dynamic Memory (DRAM) Cell

Although a variety of DRAM storage cells have been proposed over the years, a particular cell, shown in Fig. 16.18, has become the industry standard. The cell consists of a single n-channel MOSFET, known as the **access transistor**, and a **storage capacitor** C_S. The cell is appropriately known as the **one-transistor cell**.[2] The gate of the transistor is connected to the word line, and its source (drain) is connected to the bit line. Observe that only one bit line is used in DRAMs, whereas in SRAMs both the bit and $\overline{\text{bit}}$ lines are utilized. The DRAM cell stores its bit of information as charge on the cell capacitor C_S. When the cell is storing a 1, the capacitor is charged to V_{DD}; when a 0 is stored, the capacitor is discharged to zero volts.

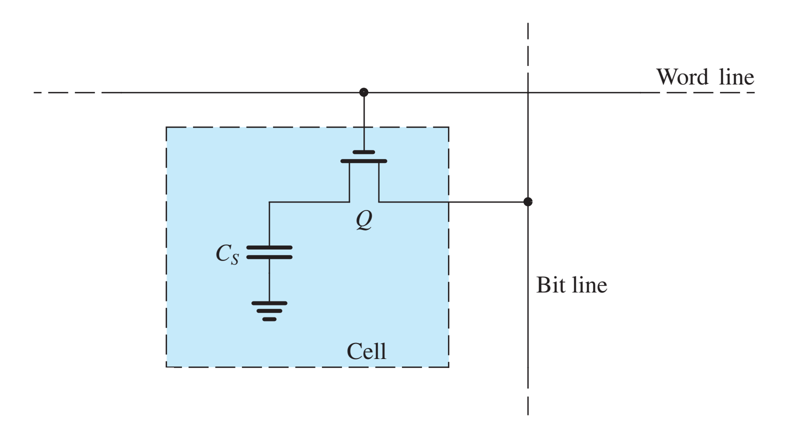


Figure 16.18 The one-transistor dynamic RAM (DRAM) cell.

[2] The name was originally used to distinguish this cell from earlier ones utilizing three transistors.



Figure 16.19 When the voltage of the selected word line is raised, the transistor conducts, thus connecting the storage capacitor C_S to the bit-line capacitor C_B.

Some explanation is needed to appreciate how the capacitor can be charged to the full supply voltage V_{DD}. Consider a write-1 operation. The word line is at V_{DD} and the bit line is at V_{DD} and the transistor is conducting, charging C_S. The transistor will cease conduction when the voltage on C_S reaches $(V_{DD} - V_t)$. This is the same problem we encountered with pass-transistor logic (PTL) in Section 15.4. The problem is overcome in DRAM design by boosting the word line to a voltage equal to $V_{DD} + V_t$. In this case the capacitor voltage for a stored 1 will be equal to the full V_{DD}. However, because of leakage effects, the capacitor charge will leak off, and hence the cell must be refreshed periodically. During **refresh**, the cell content is read and the data bit is rewritten, thus *restoring* the capacitor voltage to its proper value. Typically, the refresh operation must be performed every 5 ms to 10 ms.

Let us now consider the DRAM operation in more detail. As in the static RAM, the row decoder selects a particular row by raising the voltage of its word line. This causes all the access transistors in the selected row to become conductive, thereby connecting the storage capacitors of all the cells in the selected row to their respective bit lines. Thus the cell capacitor C_S is connected in parallel with the bit-line capacitance C_B, as indicated in Fig. 16.19. Here, it should be noted that C_S is typically 20 fF to 30 fF, whereas C_B is 10 times larger. Now, if the operation is a read, the bit line is precharged to $V_{DD}/2$. To find the change in the voltage on the bit line resulting from connecting a cell capacitor C_S to it, let the initial voltage on the cell capacitor be V_{CS} ($V_{CS} = V_{DD}$ when a 1 is stored, and $V_{CS} = 0$ V when a 0 is stored). Using charge conservation, we can write

$$C_S V_{CS} + C_B \frac{V_{DD}}{2} = (C_B + C_S)\left(\frac{V_{DD}}{2} + \Delta V\right)$$

from which we can obtain for ΔV

$$\Delta V = \frac{C_S}{C_B + C_S}\left(V_{CS} - \frac{V_{DD}}{2}\right) \tag{16.12}$$

and since $C_B \gg C_S$,

$$\Delta V \simeq \frac{C_S}{C_B}\left(V_{CS} - \frac{V_{DD}}{2}\right) \tag{16.13}$$

Now, if the cell is storing a 1, $V_{CS} = V_{DD}$, and

$$\Delta V(1) \simeq \frac{C_S}{C_B}\left(\frac{V_{DD}}{2}\right) \tag{16.14}$$

whereas if the cell is storing a 0, $V_{CS} = 0$, and

$$\Delta V(0) \simeq -\frac{C_S}{C_B}\left(\frac{V_{DD}}{2}\right) \tag{16.15}$$

Since usually C_B is much greater than C_S, these readout voltages are very small. For example, for $C_B = 10\,C_S$, $V_{DD} = 1.8$ V, $\Delta V(0)$ will be about –90 mV, and $\Delta V(1)$ will be +90 mV. This is a best-case scenario, for the 1 level in the cell might very well be below V_{DD}. Furthermore, in modern memory chips, V_{DD} is 1.2 V or even lower. In any case, we see that a stored 1 in the cell results in a small positive increment in the bit-line voltage, whereas a stored zero results in a small negative increment. Observe also that the readout process is *destructive,* since the resulting voltage across C_S will no longer be V_{DD} or 0.

The change of voltage on the bit line is detected and amplified by the column sense amplifier, causing the bit line to be driven to the full-scale value (0 or V_{DD}) of the detected signal. This amplified signal is then impressed on the storage capacitor, thus restoring its signal to the proper level (V_{DD} or 0). In this way, all the cells in the selected row are refreshed. Simultaneously, the signal at the output of the sense amplifier of the selected column is fed to the data-output line of the chip through the action of the column decoder.

The write operation proceeds similarly to the read operation, except that the data bit to be written, which is impressed on the data-input line, is applied by the column decoder to the selected bit line. Thus, if the data bit to be written is a 1, the *B*-line voltage is raised to V_{DD} (i.e., C_B is charged to V_{DD}). When the access transistor of the particular cell is turned on, its capacitor C_S will be charged to V_{DD}; thus a 1 is written in the cell. Simultaneously, all the other cells in the selected row are simply refreshed.

Although the read and write operations result in automatic refreshing of all the cells in the selected row, provision must be made for the periodic refreshing of the entire memory, typically every 5 ms to 10 ms, as specified for the particular chip. The refresh operation is carried out in a *burst mode,* one row at a time. During refresh, the chip will not be available for read or write operations. This is not a serious matter, however, since the interval required to refresh the entire chip is typically less than 2% of the time between refresh cycles. In other words, the memory chip is available for normal operation more than 98% of the time.

EXERCISES

16.7 In a particular dynamic memory chip, $C_S = 30$ fF, $C_B = 0.3$ pF, and $V_{DD} = 1.2$ V. Find the output readout voltage for a stored 1 and a stored 0. Recall that in a read operation, the bit lines are precharged to $V_{DD}/2$.
Ans. 60 mV; –60 mV

16.8 A 64-Mbit DRAM chip fabricated in a 0.4-μm CMOS technology requires 2 μm^2 per cell. If the storage array is square, estimate its dimensions. Further, if the peripheral circuitry (e.g., sense amplifiers, decoders) adds about 30% to the chip area, estimate the dimensions of the resulting chip.
Ans. 11.6 mm × 11.6 mm; 13.2 mm × 13.2 mm

16.4 Sense Amplifiers and Address Decoders

Having studied the circuits commonly used to implement the storage cells in SRAMs and DRAMs, we now consider some of the other important circuit blocks in a memory chip. The design of these circuits, commonly referred to as the **memory peripheral circuits**,

presents exciting challenges and opportunities to integrated-circuit designers: Improving the performance of peripheral circuits can result in denser and faster memory chips that dissipate less power.

16.4.1 The Sense Amplifier

Next to the storage cells, the sense amplifier is the most critical component in a memory chip. Sense amplifiers are essential to the proper operation of DRAMs, and their use in SRAMs results in speed and area improvements.

A variety of sense-amplifier designs are in use, some of which closely resemble the active-load MOS differential amplifier studied in Chapter 9. Here, we first describe a differential sense amplifier that employs positive feedback. Because the circuit is differential, it can be employed directly in SRAMs, where the SRAM cell utilizes both the B and $\overline{B}$ lines. On the other hand, the one-transistor DRAM circuit we studied in Section 16.3.2 is a single-ended circuit, utilizing one bit line only. The DRAM circuit, however, can be made to resemble a differential signal source through the use of the "dummy-cell" technique, which we shall discuss shortly. Therefore, we shall assume that the memory cell whose output is to be amplified develops a difference output voltage between the B and $\overline{B}$ lines. This signal, which can range from 20 mV to 500 mV depending on the memory type and cell design, will be applied to the input terminals of the sense amplifier. The sense amplifier in turn responds by providing a full-swing (0 to V_{DD}) signal at its output terminals. The particular amplifier circuit we shall discuss here has a rather unusual property: *Its output and input terminals are the same!*

A Sense Amplifier with Positive Feedback Figure 16.20 shows the sense amplifier together with some of the other column circuitry of a RAM chip. Note that the sense amplifier is nothing but the familiar latch formed by cross-coupling two CMOS inverters: One inverter is implemented by transistors Q_1 and Q_2, and the other by transistors Q_3 and Q_4. Transistors Q_5 and Q_6 act as switches that connect the sense amplifier to ground and V_{DD} only when data-sensing action is required. Otherwise, ϕ_s is low and the sense amplifier is turned off. This conserves power, an important consideration because usually there is one sense amplifier per column, resulting in *thousands of sense amplifiers per chip*. Note, again, that terminals x and y are both the input and the output terminals of the amplifier. As indicated, these I/O terminals are connected to the B and $\overline{B}$ lines. The amplifier is required to detect a small signal appearing between B and $\overline{B}$, and to amplify it to provide a full-swing signal at B and $\overline{B}$. For instance, if during a read operation, the cell has a stored 1, then a small positive voltage will develop between B and $\overline{B}$, with v_B higher than $v_{\overline{B}}$. The amplifier will then cause v_B to rise to V_{DD} and $v_{\overline{B}}$ to fall to 0 V. This 1 output is then directed to the chip I/O pin by the column decoder (not shown) and at the same time is used to rewrite a 1 in the DRAM cell, thus performing the restore operation that is required because the DRAM readout process is destructive.

Figure 16.20 also shows the precharge and equalization circuit. Operation of this circuit is straightforward: When ϕ_p goes high (to V_{DD}) prior to a read operation, all three transistors conduct. While Q_8 and Q_9 precharge the $\overline{B}$ and B lines to $V_{DD}/2$, transistor Q_7 helps speed up this process by equalizing the initial voltages on the two lines. This equalization is critical to the proper operation of the sense amplifier. Any voltage difference present between B and $\overline{B}$ prior to commencement of the read operation can result in erroneous interpretation by the sense amplifier of its input signal. In Fig. 16.20, we show only one of the cells in this particular column, namely, the cell whose word line is activated. The cell can be either an SRAM or a

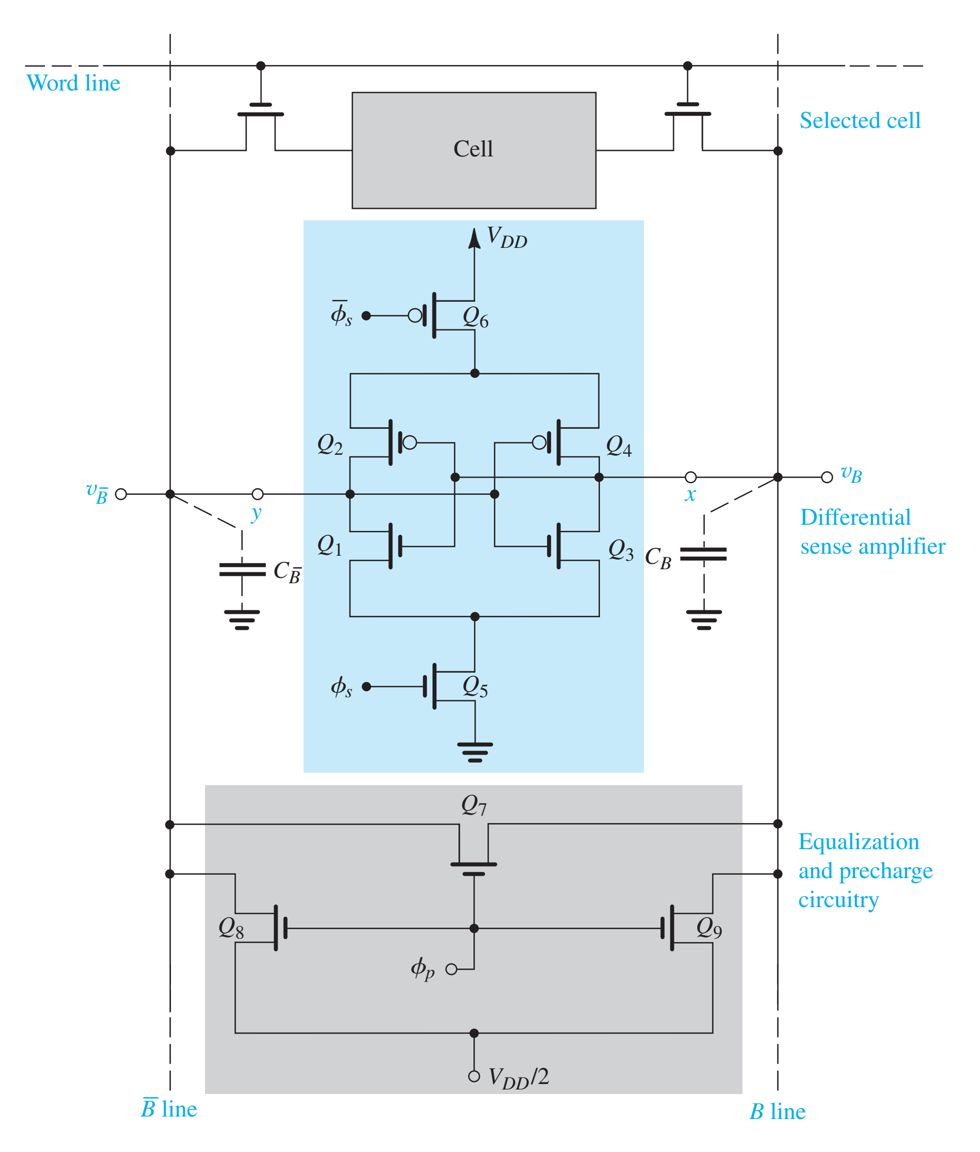


Figure 16.20 A differential sense amplifier connected to the bit lines of a particular column. This arrangement can be used directly for SRAMs (which utilize both the B and $\overline{B}$ lines). DRAMs can be turned into differential circuits by using the "dummy-cell" arrangement shown later (Fig. 16.22).

DRAM cell. All other cells in this column will not be connected to the B and $\overline{B}$ lines (because their word lines will remain low).

Let us now consider the sequence of events during a read operation:

1. The precharge and equalization circuit is activated by raising the control signal ϕ_p. This will cause the B and $\overline{B}$ lines to be at equal voltages, equal to $V_{DD}/2$. The clock ϕ_p then goes low, and the B and $\overline{B}$ lines are left to float for a brief interval.
2. The word line goes up, connecting the cell to the B and $\overline{B}$ lines. A voltage then develops between B and $\overline{B}$, with v_B higher than $v_{\overline{B}}$ if the accessed cell is storing a 1, or v_B lower than $v_{\overline{B}}$ if the cell is storing a 0. To keep the cell area small, and to

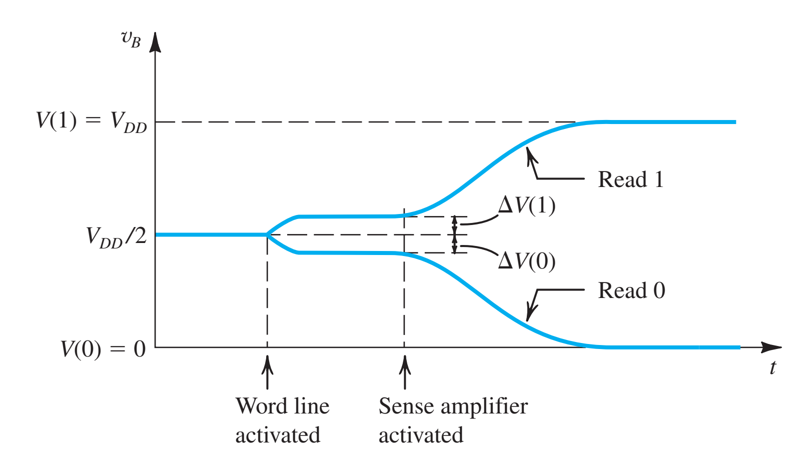


Figure 16.21 Waveforms of v_B before and after the activation of the sense amplifier. In a read-1 operation, the sense amplifier causes the initial small increment $\Delta V(1)$ to grow exponentially to V_{DD}. In a read-0 operation, the negative $\Delta V(0)$ grows to 0. Complementary signal waveforms develop on the $\overline{B}$ line.

facilitate operation at higher speeds, the readout signal, which the cell is required to provide between B and $\overline{B}$, is kept small (typically, 20–500 mV).

3. Once an adequate difference voltage signal between B and $\overline{B}$ has been developed by the storage cell, the sense amplifier is turned on by connecting it to ground and V_{DD} through Q_5 and Q_6, activated by raising the sense-control signal ϕ_s. Because initially the input terminals of the inverters are at $V_{DD}/2$, the inverters will be operating in their transition region, where the gain is high (Section 14.3). It follows that initially the latch will be operating at its unstable equilibrium point. Thus, depending on the signal between the input terminals, the latch will quickly move to one of its two stable equilibrium points (refer to the description of the latch operation in Section 16.1). This is achieved by the regenerative action, inherent in positive feedback. Figure 16.21 clearly illustrates this point by showing the waveforms of the signal on the bit line for both a read-1 and a read-0 operation. Observe that once activated, the sense amplifier causes the small initial difference, $\Delta V(1)$ or $\Delta V(0)$, provided by the cell, to grow exponentially to either V_{DD} (for a read-1 operation) or 0 (for a read-0 operation). The waveforms of the signal on the $\overline{B}$ line will be complementary to those shown in Fig. 16.21 for the B line. In the following, we quantify the process of exponential growth of v_B and $v_{\overline{B}}$.

A Closer Look at the Operation of the Sense Amplifier Developing a precise expression for the output signal of the sense amplifier shown in Fig. 16.20 is a rather complex task requiring the use of large-signal (and thus nonlinear) models of the inverter voltage-transfer characteristic, as well as taking the positive feedback into account. We will not do this here; rather, we shall consider the operation in a semiquantitative way.

Recall that at the time the sense amplifier is activated, each of its two inverters is operating in the transition region near $V_{DD}/2$. Thus, for small-signal operation, each inverter can be modeled using g_{mn} and g_{mp}, the transconductances of Q_N and Q_P, respectively, evaluated at an input bias of $V_{DD}/2$. Specifically, a small-signal v_i superimposed on $V_{DD}/2$ at the input of one of the inverters gives rise to an inverter output current signal of $(g_{mn}+g_{mp})v_i \equiv G_m v_i$. This output current is delivered to one of the capacitors, C_B or $C_{\overline{B}}$. The voltage thus developed across the capacitor is then fed back to the other inverter and is multiplied by its G_m, which gives rise

to an output current feeding the other capacitor, and so on, in a regenerative process. The positive feedback in this loop will mean that the signal around the loop, and thus v_B and $v_{\overline{B}}$, will *rise or decay exponentially* (see Fig. 16.21) with a time constant of (C_B/G_m) [or $(C_{\overline{B}}/G_m)$, since we have been assuming $C_B = C_{\overline{B}}$]. Thus, for example, in a read-1 operation we obtain

$$v_B = \frac{V_{DD}}{2} + \Delta V(1)e^{(G_m/C_B)t}, \qquad v_B \leq V_{DD} \tag{16.16}$$

whereas in a read-0 operation,

$$v_B = \frac{V_{DD}}{2} - \Delta V(0)e^{(G_m/C_B)t}, \qquad v_B \geq 0 \tag{16.17}$$

Because these expressions have been derived assuming small-signal operation, they describe the exponential growth (decay) of v_B reasonably accurately only for values close to $V_{DD}/2$. Nevertheless, they can be used to obtain a reasonable estimate of the time required to develop a particular signal level on the bit line.

Example 16.2

Consider the sense-amplifier circuit of Fig. 16.20 during the reading of a 1. Assume that the storage cell provides a voltage increment on the *B* line of $\Delta V(1) = 0.1$ V. If the NMOS devices in the amplifiers have $(W/L)_n = 0.54\ \mu\text{m}/0.18\ \mu\text{m}$ and the PMOS devices have $(W/L)_p = 2.16\ \mu\text{m}/0.18\ \mu\text{m}$, and assuming that $V_{DD} = 1.8$ V, $V_{tn} = |V_{tp}| = 0.5$ V, and $\mu_n C_{ox} = 4\,\mu_p C_{ox} = 300\ \mu\text{A/V}^2$, find the time required for v_B to reach $0.9\ V_{DD}$. Assume $C_B = 1$ pF.

Solution

First, we determine the transconductances g_{mn} and g_{mp}

$$g_{mn} = \mu_n C_{ox}\left(\frac{W}{L}\right)_n (V_{GS} - V_t)$$

$$= 300 \times \frac{0.54}{0.18}(0.9 - 0.5)$$

$$= 0.36 \text{ mA/V}$$

$$g_{mp} = \mu_p C_{ox}\left(\frac{W}{L}\right)_p (V_{GS} - |V_t|)$$

$$= 75 \times \frac{2.16}{0.18}(0.9 - 0.5) = 0.36 \text{ mA/V}$$

Thus, the inverter G_m is

$$G_m = g_{mn} + g_{mp} = 0.72 \text{ mA/V}$$

and the time constant τ for the exponential growth of v_B will be

$$\tau \equiv \frac{C_B}{G_m} = \frac{1 \times 10^{-12}}{0.72 \times 10^{-3}} = 1.4 \text{ ns}$$

Now, the time, Δt, for v_B to reach 0.9 V_{DD} can be determined from

$$0.9 \times 1.8 = 0.9 + 0.1e^{\Delta t/1.4}$$

resulting in

$$\Delta t = 2.8 \text{ ns}$$

Obtaining Differential Operation in Dynamic RAMs The sense amplifier described earlier responds to difference signals appearing between the bit lines. Thus, it is capable of rejecting interference signals that are common to both lines, such as those caused by capacitive coupling from the word lines. For this *common-mode rejection* to be effective, great care has to be taken to match both sides of the amplifier, taking into account the circuits that feed each side. This is an important consideration in any attempt to make the inherently single-ended output of the DRAM cell appear differential. We shall now discuss an ingenious scheme for accomplishing this task. Although the technique has been around for many years (see the first edition of this book, published in 1982), it is still in use today. The method is illustrated in Fig. 16.22.

Basically, each bit line is split into two identical halves. Each half-line is connected to half the cells in the column and to an additional cell, known as a *dummy cell,* having a storage

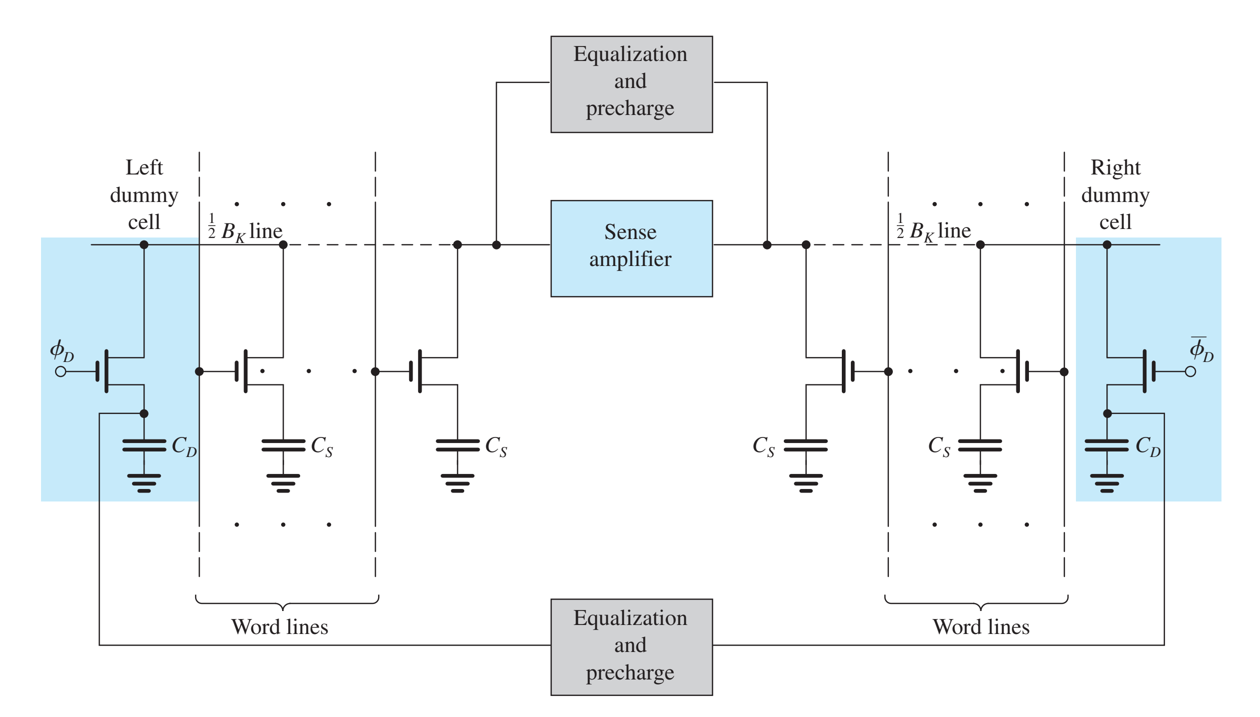


Figure 16.22 An arrangement for obtaining differential operation from the single-ended DRAM cell. Note the dummy cells at the far right and far left.

capacitor $C_D = C_S$. When a word line on the left side is selected for reading, the dummy cell on the right side (controlled by $\overline{\phi}_D$) is also selected, and vice versa; that is, when a word line on the right side is selected, the dummy cell on the left (controlled by ϕ_D) is also selected. In effect, then, the dummy cell serves as the other half of a differential DRAM cell. When the left-half bit line is in operation, the right-half bit line acts as its complement (or $\overline{B}$ line) and vice versa.

Operation of the circuit in Fig. 16.22 is as follows: The two halves of the line are precharged to $V_{DD}/2$ and their voltages are equalized. At the same time, the capacitors of the two dummy cells are precharged to $V_{DD}/2$. Then a word line is selected, and the dummy cell on the other side is enabled (with ϕ_D or $\overline{\phi}_D$ raised to V_{DD}). Thus the half-line connected to the selected cell will develop a voltage increment (around $V_{DD}/2$) of $\Delta V(1)$ or $\Delta V(0)$ depending on whether a 1 or a 0 is stored in the cell. Meanwhile, the other half of the line will have its voltage held equal to that of C_D (i.e., $V_{DD}/2$). The result is a differential signal of $\Delta V(1)$ or $\Delta V(0)$ that the sense amplifier detects and amplifies when it is enabled. As usual, by the end of the regenerative process, the amplifier will cause the voltage on one half of the line to become V_{DD} and that on the other half to become 0.

EXERCISES

16.9 It is required to reduce the time Δt of the sense-amplifier circuit in Example 16.2 by a factor of 2 by increasing g_m of the transistors (while retaining the matched design of each inverter). What must the W/L ratios of the n- and p-channel devices become?
Ans. $(W/L)_n = 6$; $(W/L)_p = 24$

16.10 If in the sense amplifier of Example 16.2, the signal available from the cell is only half as large (i.e., only 50 mV), what will Δt become?
Ans. 3.7 ns, an increase of 32%

Alternative Precharging Arrangements If it is desired to precharge the B and $\overline{B}$ lines to V_{DD}, the arrangement in Fig. 16.23(a) can be utilized. Here precharging and equalization occur when $\overline{\phi}_p$ is low. Then, just prior to the activation of the word line, $\overline{\phi}_p$ goes high. Another precharging arrangement using diode-connected NMOS transistors is shown in Fig. 16.23(b). In this case, the B and $\overline{B}$ lines are charged to $(V_{DD} - V_t)$, and Q_7 equalizes their voltages.

An Alternative Sense Amplifier Another popular implementation of the sense amplifier is the differential MOS amplifier with a current-mirror load, studied in detail in Section 9.5. Here, we present a brief overview of the operation of this versatile circuit as a sense amplifier.

The amplifier circuit is shown in Fig. 16.24 fed from the bit and $\overline{\text{bit}}$ lines (voltages v_B and $v_{\overline{B}}$). Transistors Q_1 and Q_2 are connected in the differential-pair configuration and are biased by a constant current I supplied by current source Q_5. Transistors Q_3 and Q_4 form a current mirror, which acts as the load circuit for the amplifying transistors Q_1 and Q_2. The differential nature of the amplifier aids significantly in its effectiveness as a sense amplifier: It rejects noise or interference signals that are coupled equally to the B and $\overline{B}$ lines, and amplifies only

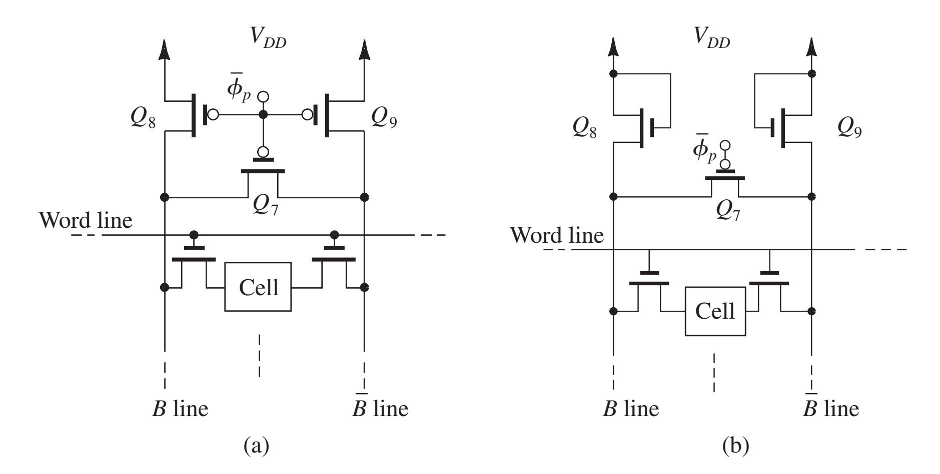


Figure 16.23 Two alternative arrangements for precharging the bit lines: **(a)** The B and $\bar{B}$ lines are precharged to V_{DD}; **(b)** the B and $\bar{B}$ lines are charged to $(V_{DD} - V_t)$.

the small difference signals that appear between B and $\bar{B}$ as a result of the read operation of a cell connected to the B and $\bar{B}$ lines.

The amplifier is designed so that in normal small-signal operation, all transistors operate in the saturation region. Figure 16.24(b) shows the amplifier in its equilibrium state with $v_B = v_{\bar{B}} = V_{DD} - V_t$. Note that we have assumed that the B and $\bar{B}$ lines are precharged to $(V_{DD} - V_t)$ using the circuit in Fig. 16.23(b). It turns out that this voltage is particularly convenient for the operation of this amplifier type as a sense amplifier. As indicated in Fig. 16.24(b), the bias current I divides equally between Q_1 and Q_2; thus each conducts a current $I/2$. The current of Q_1 is fed to the input side of the current mirror, transistor Q_3; thus the mirror provides an equal output current $I/2$ in the drain of Q_4. At the output node, we see that we have two equal and opposite currents, leaving a zero current to flow into the load capacitor. Thus, in an ideal situation of perfect matching, v_O will be equal to the voltage at the drain of Q_1.

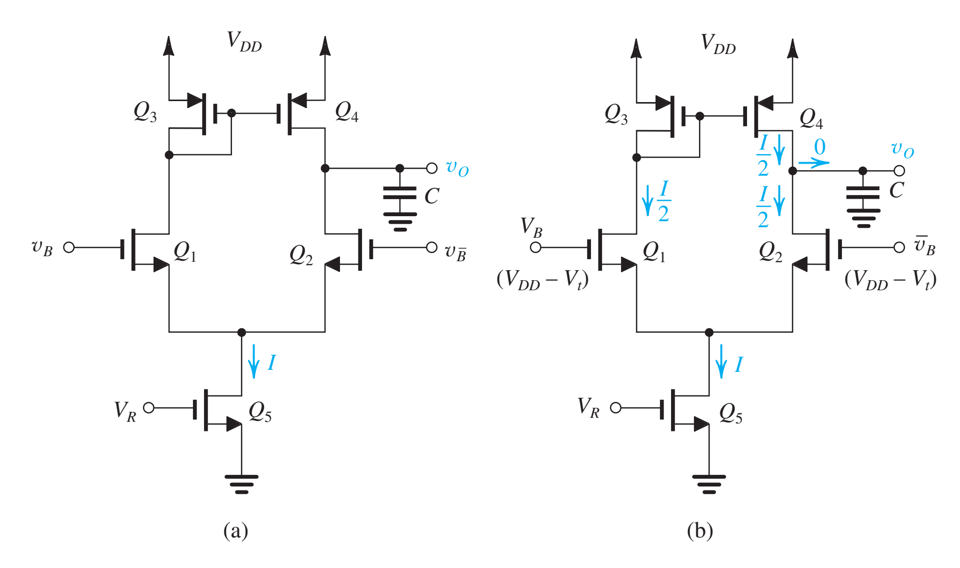


Figure 16.24 The current-mirror-loaded MOS differential amplifier as a sense amplifier.

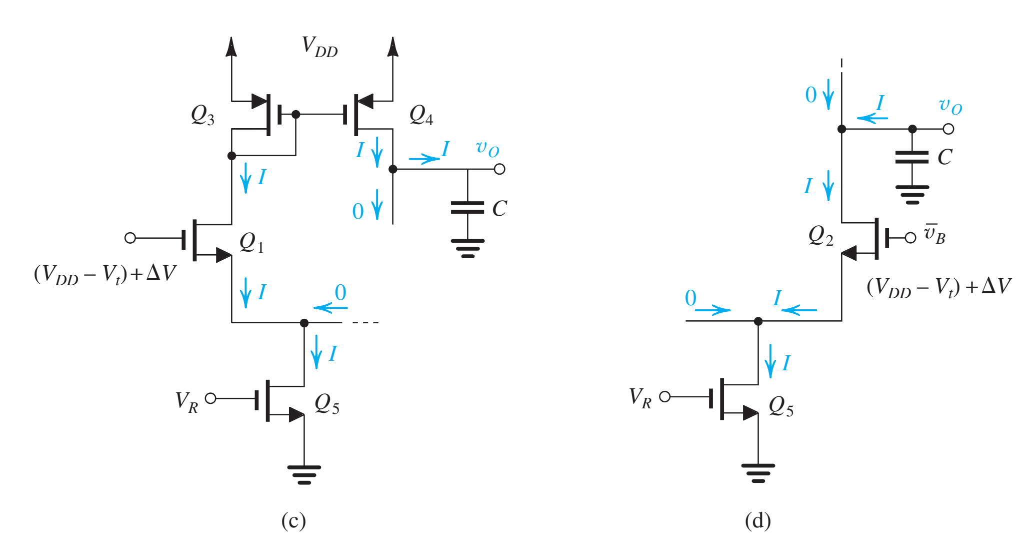


Figure 16.24 *continued*

Next consider the situation when the B line shows an incremental voltage ΔV above the voltage of the $\overline{B}$ line. As shown in Fig. 16.24(c), if ΔV is sufficiently large, Q_2 will turn off and all the bias current I will flow through Q_1 and on to Q_3. Thus the mirror output current becomes I and flows through the amplifier output terminal to the equivalent output capacitance C. Thus C will charge to V_{DD} in time Δt,

$$\Delta t = \frac{CV_{DD}}{I} \tag{16.18}$$

The complementary situation when $v_{\overline{B}}$ exceeds v_B by ΔV is illustrated in Fig. 16.24(d). Here Q_1, Q_3, and Q_4 are turned off, and Q_2 conducts all the current I. Thus capacitor C is discharged to ground by a constant current I.

An important question to answer before leaving this amplifier circuit is how large is ΔV that causes the current I to switch from one side of the differential pair to the other? The answer is given in Section 9.5 (see Fig. 9.6), namely,

$$\Delta V = \sqrt{2}V_{OV} \tag{16.19}$$

where V_{OV} is the overdrive voltage at which Q_1 and Q_2 are operating in equilibrium, that is,

$$\frac{I}{2} = \frac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)_{1,2} V_{OV}^2 \tag{16.20}$$

Finally, we note that this sense amplifier dissipates static power given by

$$P = V_{DD}I$$

Observe that increasing I reduces the time Δt in Eq. (16.18) at the expense of increased power dissipation.

EXERCISE

D16.11 It is required to design the sense amplifier in Fig. 16.24 to detect an input signal $\Delta V = 100$ mV and to provide a full output in 0.5 ns. If $C = 50$ fF and $V_{DD} = 1.8$ V, find the required current I and the power dissipation.
Ans. 180 μA; 324 μW

16.4.2 The Row-Address Decoder

As described in Section 16.2, the row-address decoder is required to select one of the 2^M word lines in response to an M-bit address input. As an example, consider the case $M = 3$ and denote the three address bits A_0, A_1, and A_2, and the eight word lines $W_0, W_1, \ldots, W_7$. Conventionally, word line W_0 will be high when $A_0 = 0$, $A_1 = 0$, and $A_2 = 0$; thus we can express W_0 as a Boolean function of A_0, A_1, and A_2,

$$W_0 = \overline{A}_0\overline{A}_1\overline{A}_2 = \overline{A_0 + A_1 + A_2}$$

Thus the selection of W_0 can be accomplished by a three-input NOR gate whose three inputs are connected to A_0, A_1, and A_2 and whose output is connected to word line 0. Word line W_3 will be high when $A_0 = 1$, $A_1 = 1$, and $A_2 = 0$; thus,

$$W_3 = A_0A_1\overline{A}_2 = \overline{\overline{A}_0 + \overline{A}_1 + A_2}$$

Thus the selection of W_3 can be realized by a three-input NOR gate whose three inputs are connected to $\overline{A}_0, \overline{A}_1$, and A_2, and whose output is connected to word line 3. We can thus see that this address decoder can be realized by eight three-input NOR gates. Each NOR gate is fed with the appropriate combination of address bits and their complements, corresponding to the word line to which its output is connected.

A simple approach to realizing these NOR functions is provided by the matrix structure shown in Fig. 16.25. The circuit shown is a dynamic one (Section 15.5). Attached to each row line is a p-channel device that is activated, prior to the decoding process, using the precharge control signal ϕ_P. During precharge (ϕ_P low), all the word lines are pulled high to V_{DD}. It is assumed that at this time the address input bits have not yet been applied and all the inputs are low; hence there is no need for the circuit to include the evaluation transistor utilized in dynamic logic gates. Then, the decoding operation begins when the address bits and their complements are applied. Observe that the NMOS transistors are placed so that the word lines not selected will be discharged. For any input combination, only one word line will not be discharged, and thus its voltage remains high at V_{DD}. For instance, row 0 will be high only when $A_0 = 0$, $A_1 = 0$, and $A_2 = 0$; this is the only combination that will result in all three transistors connected to row 0 being cut off. Similarly, row 3 has transistors connected to $\overline{A}_0, \overline{A}_1$, and A_2, and thus it will be high when $A_0 = 1$, $A_1 = 1$, $A_2 = 0$, and so on. After the decoder outputs have stabilized, the output lines are connected to the word lines of the array, usually via clock-controlled transmission gates. This decoder is known as a NOR decoder. Observe that because of the precharge operation, the decoder circuit does not dissipate static power.

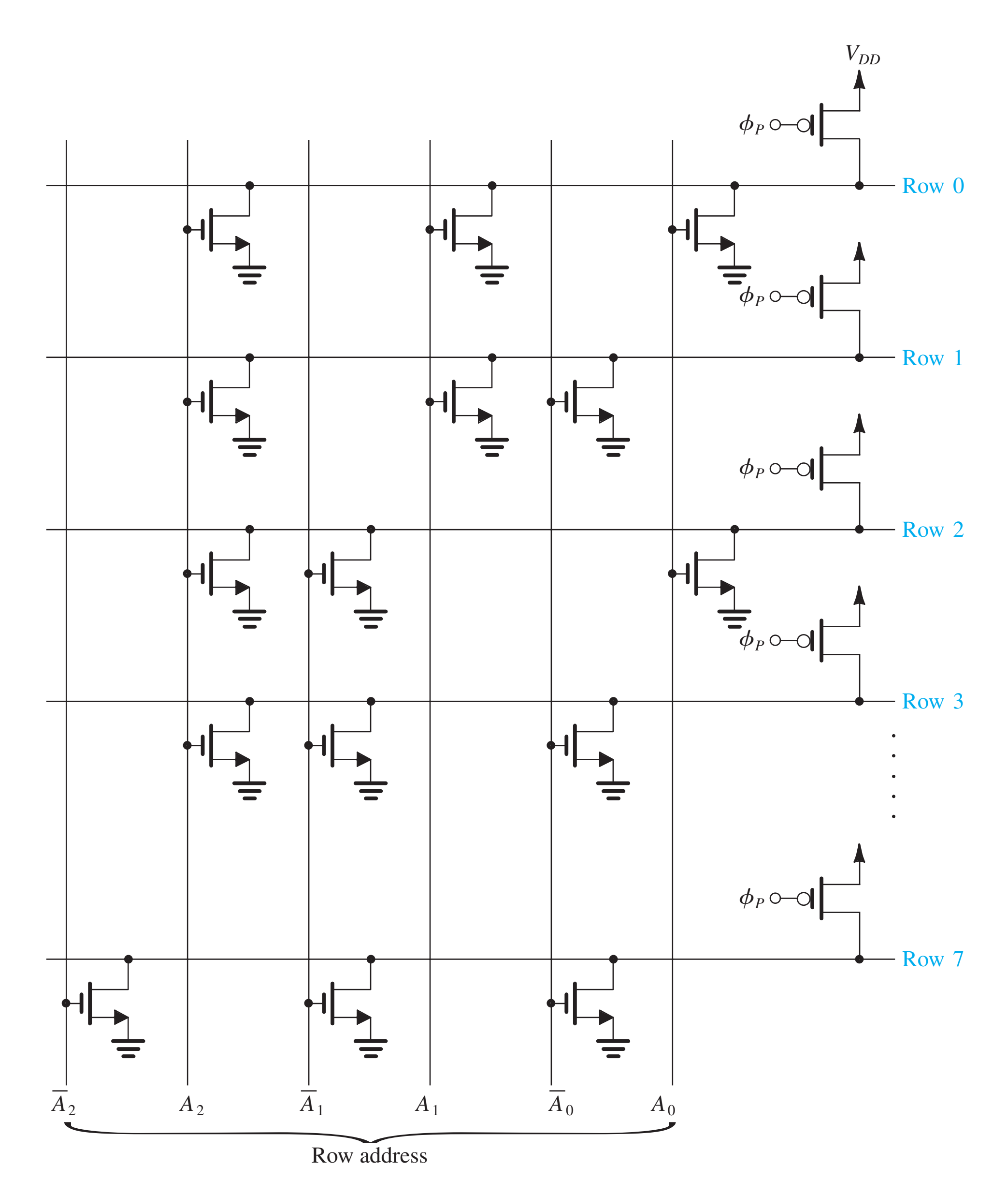


Figure 16.25 A NOR address decoder in array form. One out of eight lines (row lines) is selected using a 3-bit address.

EXERCISE

16.12 How many transistors are needed for a NOR row decoder with an M-bit address?

Ans. $M2^M$ NMOS $+ 2^M$ PMOS $= 2^M(M+1)$

16.4.3 The Column-Address Decoder

From the description in Section 16.2, the function of the column-address decoder is to connect one of the 2^N bit lines to the data I/O line of the chip. As such, it is a multiplexer and can be implemented using pass-transistor logic (Section 15.4) as shown in Fig. 16.26. Here, each bit line is connected to the data I/O line through an NMOS transistor. The gates of the pass transistors are controlled by 2^N lines, one of which is selected by a NOR decoder similar to

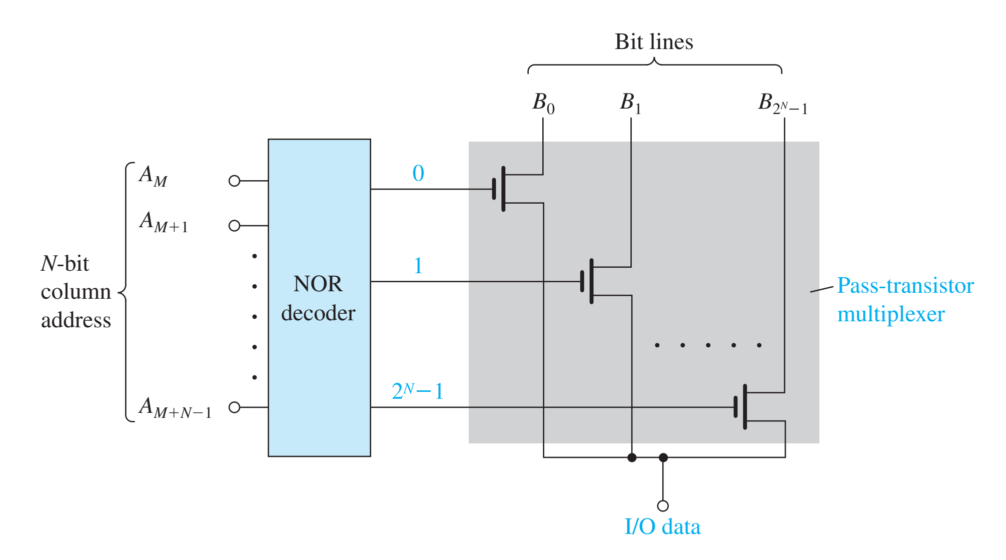


Figure 16.26 A column decoder realized by a combination of a NOR decoder and a pass-transistor multiplexer.

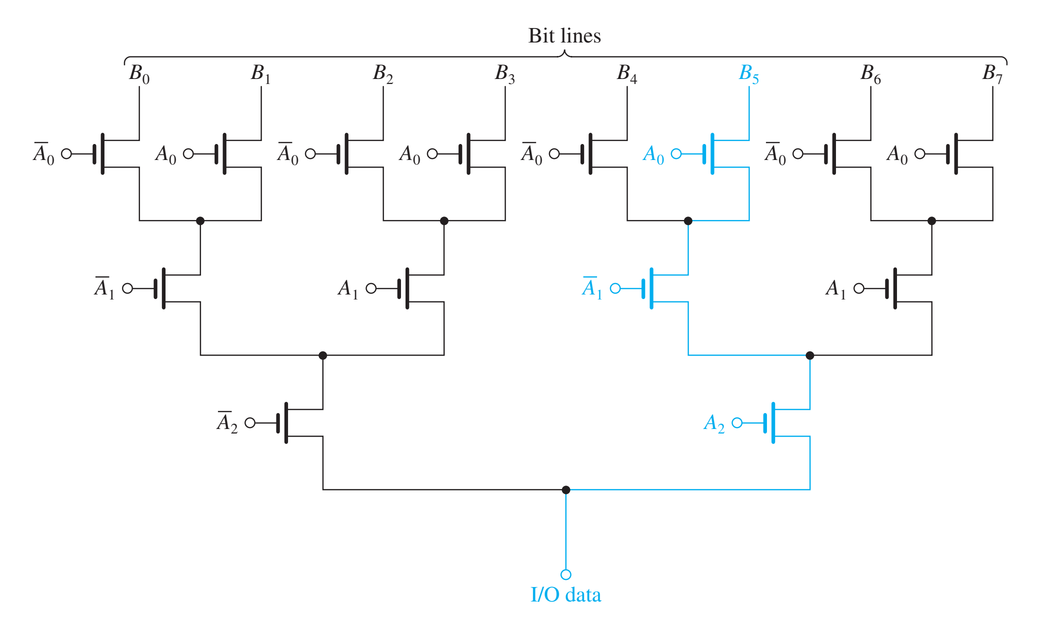


Figure 16.27 A tree column decoder. Note that the colored path shows the transistors that are conducting when $A_0 = 1$, $A_1 = 0$, and $A_2 = 1$, the address that results in connecting B_5 to the data line.

that used for decoding the row address. Finally, note that better performance can be obtained by utilizing transmission gates in place of NMOS transistors (Section 15.4). In such a case, however, the decoder needs to provide complementary output signals.

An alternative implementation of the column decoder that uses a smaller number of transistors (but at the expense of slower speed of operation) is shown in Fig. 16.27. This circuit, known as a *tree decoder,* has a simple structure of pass transistors. Unfortunately, since a relatively large number of transistors can exist in the signal path, the resistance of the bit lines increases, and the speed decreases correspondingly.

EXERCISE

16.13 How many transistors are needed for a tree decoder when there are 2^N bit lines?
Ans. $2(2^N - 1)$

16.4.4 Pulse-Generation Circuits

Memory chips require a large number of pulse signals, sometimes with intricate timing relationships among them. It is not our purpose here to study this important subject; rather, we present two simple circuits that find widespread applicability in memory-chip timing as well as in other digital system components, such as microprocessors.

The Ring Oscillator The ring oscillator is formed by connecting an *odd* number of inverters in a loop. Although usually at least five inverters are used, we illustrate the principle of operation using a ring of three inverters, as shown in Fig. 16.28(a). Figure 16.28(b) shows the waveforms obtained at the outputs of the three inverters. These waveforms are idealized in the sense that their edges have zero rise and fall times. Nevertheless, they will serve to explain the circuit operation.

Observe that a rising edge at node 1 propagates through gates 1, 2, and 3 to return inverted after a delay of $3t_P$. This falling edge then propagates, and returns with the original (rising) polarity after another $3t_P$ interval. It follows that the circuit oscillates with a period of $6t_P$ or correspondingly with frequency $1/6t_P$. In general, a ring with N inverters (where N must be odd) will oscillate with a period of $2Nt_P$ and frequency $1/2Nt_P$.

As a final remark, we note that the ring oscillator provides a relatively simple means for measuring the inverter propagation delay.

EXERCISE

16.14 Find the frequency of oscillation of a ring of five inverters if the inverter propagation delay is specified to be 1 ns.
Ans. 100 MHz

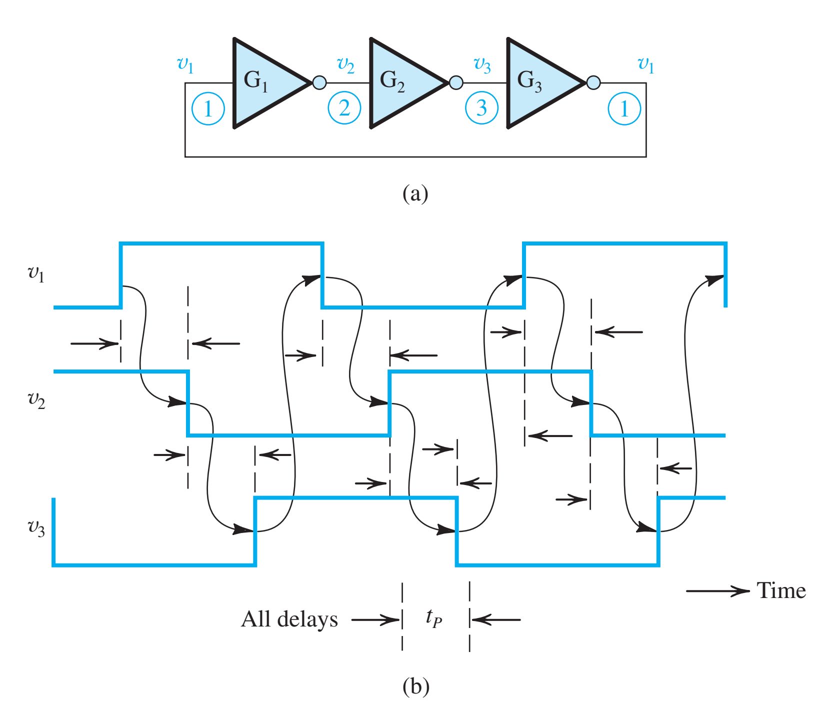


Figure 16.28 **(a)** A ring oscillator formed by connecting three inverters in cascade. (Normally at least five inverters are used.) **(b)** The resulting waveform. Observe that the circuit oscillates with frequency $1/6t_P$.

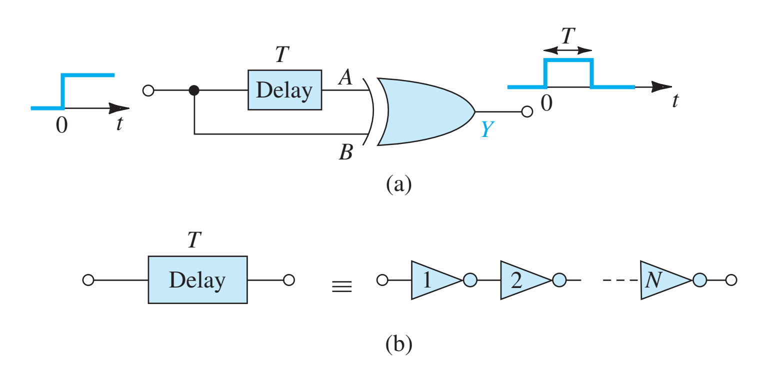


Figure 16.29 **(a)** A one-shot or monostable circuit. Utilizing a delay circuit with a delay T and an XOR gate, this circuit provides an output pulse of width T. **(b)** The delay circuit can be implemented as the cascade of N inverters where N is even, in which case $T = Nt_P$.

A One-Shot or Monostable Multivibrator Circuit The one-shot or monostable multivibrator circuit provides, when triggered, a single output pulse with a predetermined width.[3] A variety of circuits exist for implementing the one-shot function, and some using op amps will be studied in Section 18.6. Here, in Fig. 16.29(a), we show a circuit commonly

[3]The name "monostable" arises because this class of circuits has one stable state, which is the quiescent state. When a trigger is applied, the circuit moves to its quasi-stable state and stays in it for a predetermined length of time (the width of the output pulse). It then switches back automatically to the stable state.

used in digital IC design. The circuit utilizes an exclusive-OR (XOR) gate together with a delay circuit. Recalling that the XOR gate provides a high output only when its two inputs are dissimilar, we see that prior to the arrival of the input positive step, the output will be low. When the input goes high, only the *B* input of the XOR will be high and thus its output will go high. The high input will reach input *A* of the XOR *T* seconds later, at which time both inputs of the XOR will be high and thus its output will go low. We thus see that the circuit produces an output pulse with a duration *T* equal to the delay of the delay block for each transition of the input signal. The delay block can be implemented by connecting an even number of inverters in cascade as shown in Fig. 16.29(b).

16.5 Read-Only Memory (ROM)

As mentioned in Section 16.2, read-only memory (ROM) is memory that contains fixed data patterns. It is used in a variety of digital system applications. Currently, a very popular application is the use of ROM in microprocessor systems to store the instructions of the system's basic operating program. ROM is particularly suited for such an application because it is nonvolatile; that is, it retains its contents when the power supply is switched off.

A ROM can be viewed as a combinational logic circuit for which the input is the collection of address bits of the ROM and the output is the set of data bits retrieved from the addressed location. This viewpoint leads to the application of ROMs in code conversion—that is, in changing the code of the signal from one system (say, binary) to another. Code conversion is employed, for instance, in secure communication systems, where the process is known as *scrambling*. It consists of feeding the code of the data to be transmitted to a ROM that provides corresponding bits in a (supposedly) secret code. The reverse process, which also uses a ROM, is applied at the receiving end.

In this section we will study various types of read-only memory. These include fixed ROM, which we refer to simply as ROM, programmable ROM (PROM), erasable programmable ROM (EPROM), and flash memory.

16.5.1 A MOS ROM

Figure 16.30 shows a simplified 32-bit (or 8-word × 4-bit) MOS ROM. As indicated, the memory consists of an array of *n*-channel MOSFETs whose gates are connected to the word lines, whose sources are grounded, and whose drains are connected to the bit lines. Each bit line is connected to the power supply via a PMOS load transistor, in the manner of pseudo-NMOS logic (Section 15.3). An NMOS transistor exists in a particular cell if the cell is storing a 0; a cell storing a 1 has no MOSFET. This ROM can be thought of as 8 words of 4 bits each. The row decoder selects one of the 8 words by raising the voltage of the corresponding word line. The cell transistors connected to this word line will then conduct, thus pulling the voltage of the bit lines (to which transistors in the selected row are connected) down from V_{DD} to a voltage close to ground voltage (the logic-0 level). The bit lines that correspond to cells (of the selected word) without transistors (i.e., the cells that are storing a logic 1) will remain at the power-supply voltage (logic 1) because of the action of the pull-up PMOS load devices. In this way, the bits of the addressed word can be read.

A disadvantage of the ROM circuit in Fig. 16.30 is that it dissipates static power. Specifically, when a word is selected, the transistors in this particular row will conduct static current that is supplied by the PMOS load transistors. Static power dissipation can

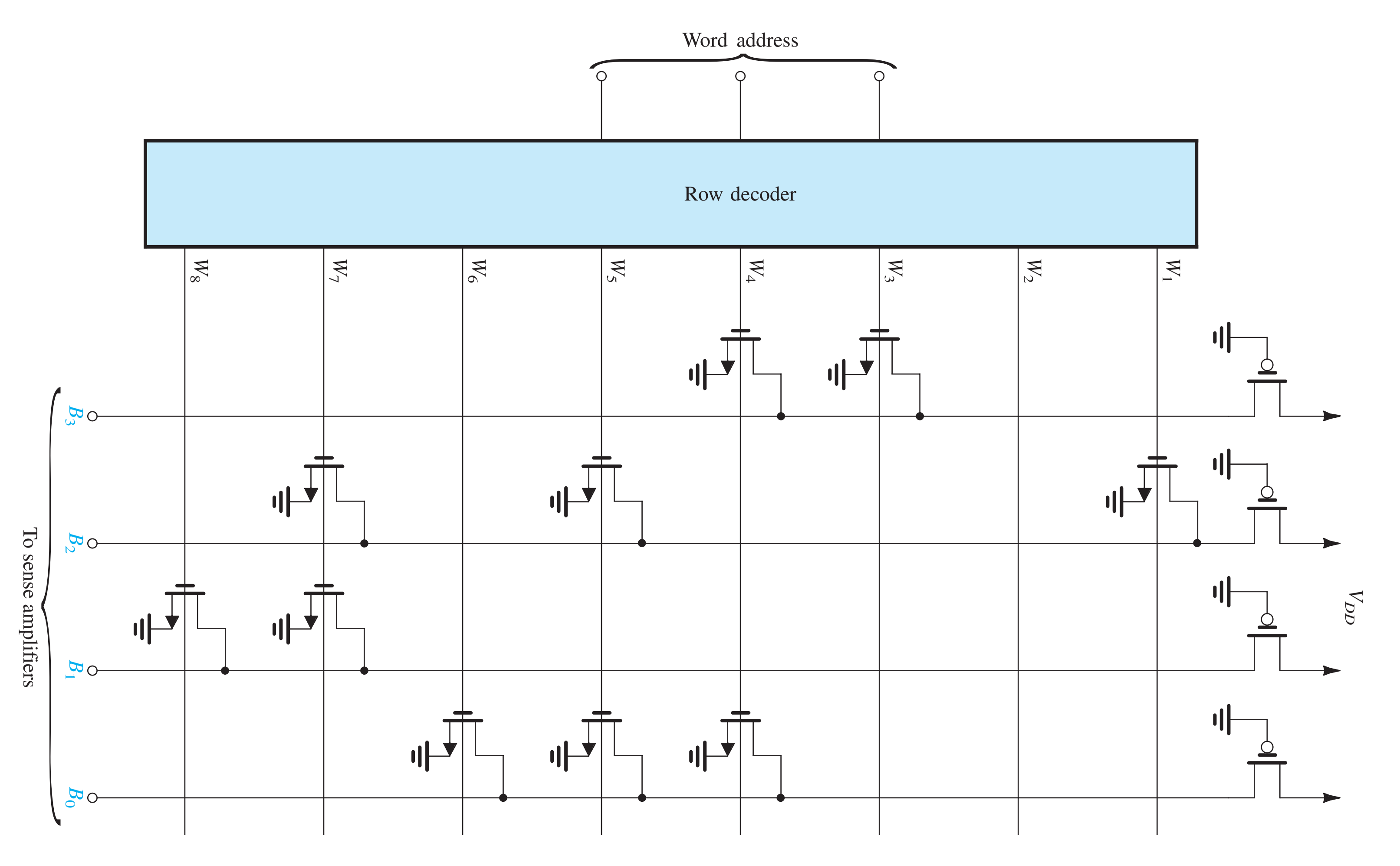


Figure 16.30 A simple MOS ROM organized as 8 words × 4 bits.

be eliminated by a simple change. Rather than grounding the gate terminals of the PMOS transistors, we can connect them to a precharge line ϕ that is normally high. Just before a read operation, ϕ is lowered and the bit lines are precharged to V_{DD} through the PMOS transistors. The precharge signal ϕ then goes high, and the word line is selected. The bit lines that have transistors in the selected word are then discharged, thus indicating stored zeros, whereas those lines for which no transistor is present remain at V_{DD}, indicating stored ones.

EXERCISE

16.15 The purpose of this exercise is to estimate the various delay times involved in the operation of a ROM. Consider the ROM in Fig. 16.30 with the gates of the PMOS devices disconnected from ground and connected to a precharge control signal ϕ. Let all the NMOS devices have $W/L = 6$ μm/2 μm and all the PMOS devices have $W/L = 24$ μm/2 μm. Assume that $\mu_n C_{ox} = 50$ μA/V^2, $\mu_p C_{ox} = 20$ μA/V^2, $V_{tn} = -V_{tp} = 1$ V, and $V_{DD} = 5$ V.

(a) During the precharge interval, ϕ is lowered to 0 V. Estimate the time required to charge a bit line from 0 V to 5 V. Use, as an average charging current, the current supplied by a PMOS transistor at a bit-line voltage halfway through the 0-V to 5-V excursion (i.e., 2.5 V). The bit-line capacitance is 2 pF. Note that all NMOS transistors are cut off at this time.

(b) After completion of the precharge interval and the return of ϕ to V_{DD}, the row decoder raises the voltage of the selected word line. Because of the finite resistance and capacitance of the word line, the voltage rises exponentially toward V_{DD}. If the resistance of each of the polysilicon word lines is 3 kΩ and the capacitance between the word line and ground is 3 pF, what is the (10% to 90%) rise time of the word-line voltage? What is the voltage reached at the end of one time constant?

(c) We account for the exponential rise of the word-line voltage by approximating the word-line voltage by a step equal to the voltage reached in one time constant. Find the interval Δt required for an NMOS transistor to discharge the bit line and lower its voltage by 0.5 V. (It is assumed that the sense amplifier needs a 0.5-V change at its input to detect a low bit value.)

Ans. (a) 6.1 ns; (b) 19.8 ns, 3.16 V; (c) 2.9 ns

16.5.2 Mask-Programmable ROMs

The data stored in the ROMs discussed thus far is determined at the time of fabrication, according to the user's specifications. However, to avoid having to custom-design each ROM from scratch (which would be extremely costly), ROMs are manufactured using a process known as **mask programming**. As explained in Appendix A, integrated circuits are fabricated on a wafer of silicon using a sequence of processing steps that include photomasking, etching, and diffusion. In this way, a pattern of junctions and interconnections is created on the surface of the wafer. One of the final steps in the fabrication process consists of coating the surface of the wafer with a layer of aluminum and then selectively (using a mask) etching away portions of the aluminum, leaving aluminum only where interconnections are desired. This last step can be used to program (i.e., to store a desired pattern in) a ROM. For instance, if the ROM is made of MOS transistors as in Fig. 16.30, MOSFETs can be included at all bit locations, but only the gates of those transistors where 0s are to be stored are connected to the word lines;

the gates of transistors where 1s are to be stored are not connected. This pattern is determined by the mask, which is produced according to the user's specifications.

The economic advantages of the mask programming process should be obvious: All ROMs are fabricated similarly; customization occurs only during one of the final steps in fabrication.

16.5.3 Programmable ROMs (PROMs, EPROMs, and Flash)

PROMs are ROMs that can be programmed by the user, but only once. A typical arrangement employed in BJT PROMs involves using polysilicon fuses to connect the emitter of each BJT to the corresponding digit line. Depending on the desired content of a ROM cell, the fuse can be either left intact or blown out using a large current. The programming process is obviously irreversible.

An erasable programmable ROM, or EPROM, is a ROM that can be erased and reprogrammed as many times as the user wishes. It is therefore the most versatile type of read-only memory. It should be noted, however, that the process of erasure and reprogramming is time consuming and is intended to be performed only infrequently.

State-of-the-art EPROMs use variants of the memory cell whose cross section is shown in Fig. 16.31(a). The cell is basically an enhancement-type *n*-channel MOSFET with two gates made of polysilicon material.[4] One of the gates is not electrically connected to any other part of the circuit; rather, it is left floating and is appropriately called a **floating gate**. The other gate, called a **select gate**, functions in the same manner as the gate of a regular enhancement MOSFET.

The MOS transistor of Fig. 16.31(a) is known as a **floating-gate transistor** and is given the circuit symbol shown in Fig. 16.31(b). In this symbol the broken line denotes the floating gate. The memory cell is known as the **stacked-gate cell**.

Let us now examine the operation of the floating-gate transistor. Before the cell is programmed (we will shortly explain what this means), no charge exists on the floating gate and the device operates as a regular *n*-channel enhancement MOSFET. It thus exhibits the i_D–v_{GS} characteristic shown as curve (*a*) in Fig. 16.32. Note that in this case the threshold voltage (V_t) is rather low. This state of the transistor is known as the **not-programmed state**.

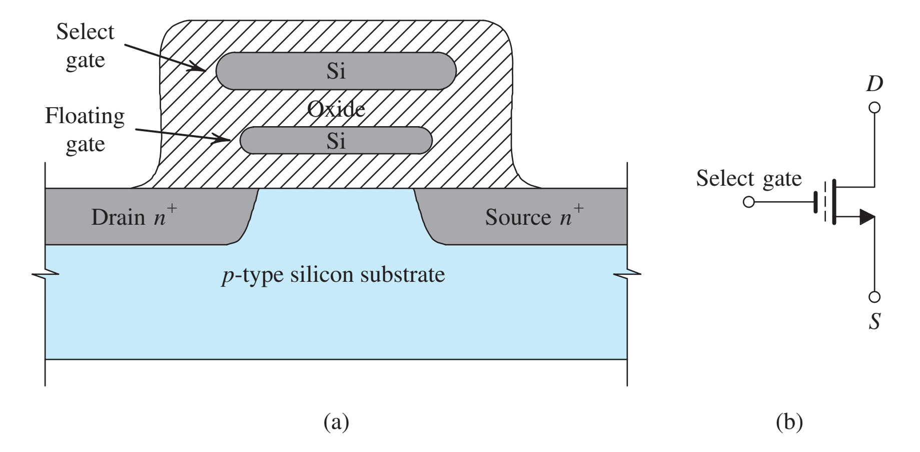


Figure 16.31 **(a)** Cross section and **(b)** circuit symbol of the floating-gate transistor used as an EPROM cell.

[4] See Appendix A for a description of silicon-gate technology.

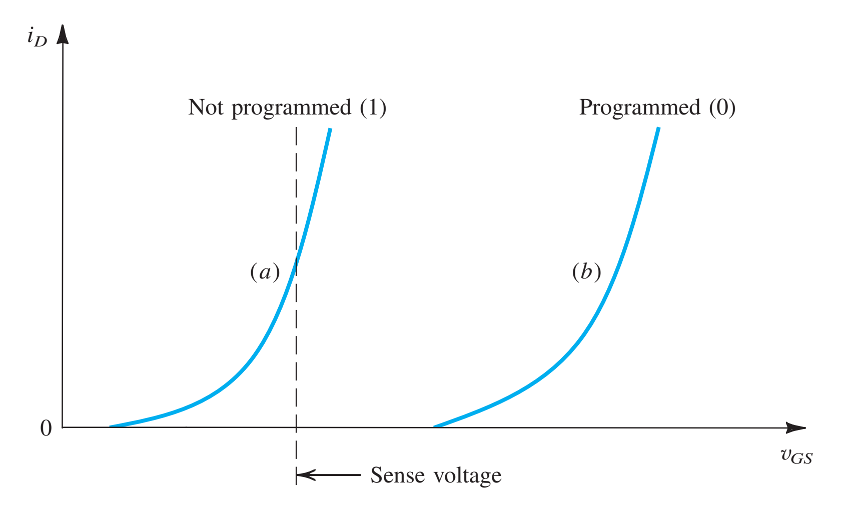


Figure 16.32 Illustrating the shift in the i_D–v_{GS} characteristic of a floating-gate transistor as a result of programming.

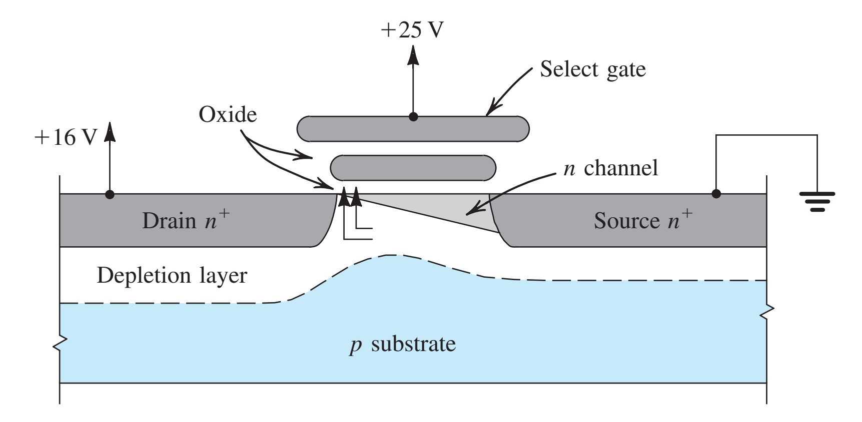


Figure 16.33 The floating-gate transistor during programming.

It is one of two states in which the floating-gate transistor can exist. Let us arbitrarily take the not-programmed state to represent a stored 1. That is, a floating-gate transistor whose i_D–v_{GS} characteristic is that shown as curve (*a*) in Fig. 16.32 will be said to be storing a 1.

To program the floating-gate transistor, a large voltage (16–20 V) is applied between its drain and source. Simultaneously, a large voltage (about 25 V) is applied to its select gate. Figure 16.33 shows the floating-gate MOSFET during programming. In the absence of any charge on the floating gate, the device behaves as a regular *n*-channel enhancement MOSFET: An *n*-type inversion layer (channel) is created at the wafer surface as a result of the large positive voltage applied to the select gate. Because of the large positive voltage at the drain, the channel has a tapered shape.

The drain-to-source voltage accelerates electrons through the channel. As these electrons reach the drain end of the channel, they acquire high kinetic energy and are referred to as *hot electrons*. The large positive voltage on the select gate (greater than the drain voltage) establishes an electric field in the insulating oxide. This electric field attracts the hot electrons and accelerates them (through the oxide) toward the floating gate. In this way the floating gate is charged, and the charge that accumulates on it becomes trapped.

Fortunately, the process of charging the floating gate is self-limiting. The negative charge that accumulates on the floating gate reduces the strength of the electric field in the oxide to the point that it eventually becomes incapable of accelerating any more of the hot electrons.

Let us now inquire about the effect of the floating gate's negative charge on the operation of the transistor. The negative charge trapped on the floating gate will cause electrons to be repelled from the surface of the substrate. This implies that to form a channel, the positive voltage that has to be applied to the select gate will have to be greater than that required when the floating gate is not charged. In other words, the threshold voltage V_t of the programmed transistor will be higher than that of the not-programmed device. In fact, programming causes the i_D–v_{GS} characteristic to shift to the curve labeled (*b*) in Fig. 16.32. In this state, known as the *programmed state,* the cell is said to be storing a 0.

Once programmed, the floating-gate device retains its shifted i–v characteristic (curve *b*) even when the power supply is turned off. In fact, extrapolated experimental results indicate that the device can remain in the programmed state for as long as 100 years!

Reading the content of the stacked-gate cell is easy: A voltage V_{GS} somewhere between the low and high threshold values (see Fig. 16.32) is applied to the selected gate. While a programmed device (one that is storing a 0) will not conduct, a not-programmed device (one that is storing a 1) will conduct heavily.

To return the floating-gate MOSFET to its not-programmed state, the charge stored on the floating gate has to be returned to the substrate. This *erasure* process can be accomplished by illuminating the cell with ultraviolet light of the correct wavelength (2537 Å) for a specified duration. The ultraviolet light imparts sufficient photon energy to the trapped electrons to allow them to overcome the inherent energy barrier, and thus be transported through the oxide, back to the substrate. To allow this erasure process, the EPROM package contains a quartz window. Finally, it should be noted that the device is extremely durable, and can be erased and programmed many times.

A more versatile programmable ROM is the electrically erasable PROM (or EEPROM). As the name implies, an EEPROM can be erased and reprogrammed electrically without the need for ultraviolet illumination. EEPROMs utilize a variant of the floating-gate MOSFET. An important class of EEPROMs using a floating-gate variant and implementing block erasure are referred to as **flash memories**. The name "flash" arises because many rows can be erased "in a flash," certainly very rapidly in comparison to the lengthy process of erasing by means of ultraviolet light. Flash memories have virtually replaced the EPROM variety and are currently (2014) very popular.

16.6 CMOS Image Sensors

We conclude this chapter by presenting a very important functional block whose overall structure is very similar to that of a memory array: The CMOS image sensor is the basic image-capturing element in digital cameras (including smartphone cameras).

An image consists of a two-dimensional array of *pixels*, where each pixel indicates light intensity at its location in the array. A CMOS image sensor consists of a two-dimensional array of *pixel circuits*, where each pixel circuit measures light intensity and is usually a few square microns in size. Pixel circuits are accessed through a set of horizontal row-access lines, analogous to the word lines in memory arrays, and intensities are read out through a set of vertical lines, analogous to the bit lines in a memory array. However, here the vertical lines carry analog signals.

A pixel circuit, called an *active pixel sensor* (APS), is shown in Fig. 16.34. Prior to an image capture, transistor Q_P resets the interval node X to a high voltage. The photodiode D

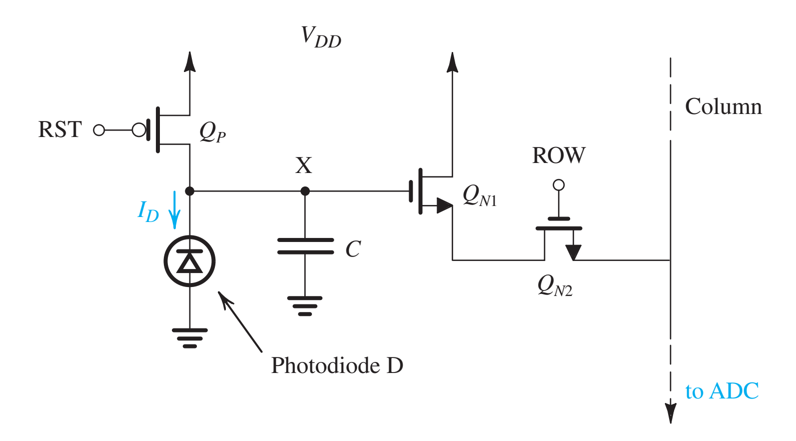


Figure 16.34 Pixel circuit in a CMOS image sensor.

is thus reverse biased, and its current I_D is essentially proportional to light intensity. Over the sensing interval T, the discharge of the parasitic capacitor C by current I_D causes a voltage drop ΔV to occur at X. This voltage change is then read out onto the column line by activating a source follower Q_{N1} and a current source (not shown, connected to the column line) and a switch Q_{N2}. The resulting analog signal on the column line is then fed to an analog-to-digital converter (ADC) to provide a digital number corresponding to the light intensity of this pixel. The digital data thus produced can be used for further digital processing of the captured image.

BLINDING FLASH:

Since its invention in 1980 by Toshiba, flash memory based on the floating-gate MOS transistor has expanded into every possible field of computing. Because of its nonvolatility, flash memory has become largely responsible for the dominance of mobile digital devices. Increasingly, flash-based solid-state drives (SSDs) are overtaking hard disk drives in enterprise memory systems. SSDs of more than 1 terabyte with no moving parts are becoming available for a few hundred dollars in technologies as small as 20 nm. In many applications, high data rates of up to 12 Gb/s allow total replacement of volatile DRAMs in handheld devices. At the other end of the scale, flash-filled USB drives with gigabyte capacities have effectively replaced the need for DVDs in today's laptops.

Summary

- Flip-flops employ one or more latches. The basic static latch is a bistable circuit implemented using two inverters connected in a positive-feedback loop. The latch can remain in either stable state indefinitely.
- As an alternative to the positive-feedback approach, memory can be provided through the use of charge storage. A number of CMOS flip-flops are realized this way, including some master–slave D flip-flops.
- A random-access memory (RAM) is one in which the time required for storing (writing) information and for retrieving (reading) information is independent of the physical location (within the memory) at which the information is stored.
- The major part of a memory chip consists of the cells in which the bits are stored and that are typically organized in a square matrix. A cell is selected for reading or writing by

activating its row, via the row-address decoder, and its column, via the column-address decoder. The sense amplifier detects the content of the selected cell and provides a full-swing version of it to the data-output terminal of the chip.

- There are two kinds of MOS RAM: static and dynamic. Static RAMs (SRAMs) employ flip-flops as the storage cells. In a dynamic RAM (DRAM), data is stored on a capacitor and thus must be periodically refreshed. DRAM chips provide the highest possible storage capacity for a given chip area.
- Two circuits have emerged as the near-universal choice in implementing the storage cell: the six-transistor SRAM cell and the one-transistor DRAM cell.
- Although sense amplifiers are utilized in SRAMs to speed up operation, they are essential in DRAMs. A particular type of sense amplifier is a differential circuit that employs positive feedback to obtain an output signal that grows exponentially toward either V_{DD} or 0.
- Read-only memory (ROM) contains fixed data patterns that are stored at the time of fabrication and cannot be changed by the user. On the other hand, the contents of an erasable programmable ROM (EPROM) can be changed by the user. The erasure and reprogramming is a time-consuming process and is performed only infrequently.
- Some EPROMS utilize floating-gate MOSFETs as the storage cells. The cell is programmed by applying (to the selected gate) a high voltage, which in effect changes the threshold voltage of the MOSFET. Erasure is achieved by illuminating the chip by ultraviolet light. Even more versatile, EEPROMs can be erased and reprogrammed electrically. These are called flash memories and are currently in widespread use.
- CMOS image sensors are organized in arrays very similar to those used in memories. Each pixel circuit measures the light intensity at its pixel and provides this information on its column line in analog form, which is converted into a digital signal by means of an analog-to-digital converter (ADC).

PROBLEMS

Section 16.1: Latches and Flip-Flops

16.1 Consider the latch of Fig. 16.1 with the two inverters identical and each characterized by $V_{OL} = 0$ V, $V_{OH} = 5$ V, $V_{IL} = 2$ V, and $V_{IH} = 3$ V. Let the transfer characteristic of each inverter be approximated by three straight-line segments. Sketch the transfer characteristic of the feedback loop of the latch and give the coordinates of points A, B, and C [refer to Fig. 16.1(b)]. What is the gain at C? What is the width of the transition region?

D 16.2 Sketch the standard CMOS circuit implementation of the SR flip-flop shown in Fig. 16.3.

D 16.3 Sketch the logic-gate implementation of an SR flip-flop utilizing two cross-coupled NAND gates. Clearly label the output terminals and the input trigger terminals. Provide the truth table and describe the operation.

D 16.4 For the SR flip-flop of Fig. 16.4, show that if each of the two inverters utilizes matched transistors, that is, $(W/L)_p = (\mu_n/\mu_p)(W/L)_n$, then the minimum W/L that each of Q_5–Q_8 must have so that switching occurs is $2(W/L)_n$. Give the sizes of all eight transistors if the flip-flop is fabricated in a 0.13-μm process for which $\mu_n = 4\mu_p$. Use the minimum channel length for all transistors and the minimum size ($W/L = 1$) for Q_1 and Q_3.

D 16.5 Repeat part (a) of the problem in Example 16.1 for the case of inverters that do not use matched Q_N and Q_P. Rather, assume that each of the inverters uses $(W/L)_p = (W/L)_n =$ 0.27 μm/0.18 μm. Find the threshold voltage of each inverter. Then determine the value required for the W/L of each of Q_5 to Q_8 so that the flip-flop switches. (*Hint:* Refer to Table 14.2.)

D 16.6 In this problem we investigate the effect of velocity saturation (Section 15.1.3) on the design of the SR flip-flop in Example 16.1. Specifically, answer part (a) of the question

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

in Example 16.1, taking into account the fact that for this technology, V_{DSsat} for n-channel devices is 0.6 V and $|V_{DSsat}|$ for p-channel devices is 1 V. Assume $\lambda_n = |\lambda_p| = 0.1\ \text{V}^{-1}$. What is the minimum required value for $(W/L)_5$ and for $(W/L)_6$? Comment on this value relative to that found in Example 16.1. (*Hint*: Refer to Eq. 15.11.)

D 16.7 The CMOS SR flip-flop in Fig. 16.4 is fabricated in a 0.13-μm process for which $\mu_n C_{ox} = 4\mu_p C_{ox} = 500\ \mu\text{A/V}^2$, $V_{tn} = |V_{tp}| = 0.4$ V, and $V_{DD} = 1.2$ V. The inverters have $(W/L)_n = 0.2\ \mu\text{m}/0.13\ \mu\text{m}$ and $(W/L)_p = 0.8\ \mu\text{m}/0.13\ \mu\text{m}$. The four NMOS transistors in the set–reset circuit have equal W/L ratios.

(a) Determine the minimum value required for this ratio to ensure that the flip-flop will switch.
(b) If a ratio twice the minimum is selected, determine the minimum required width of the set and reset pulses to ensure switching. Assume that the total capacitance between each of the Q and $\overline{Q}$ nodes and ground is 15 fF.

D 16.8 The clocked SR flip-flop in Fig. 16.4 is not a fully complementary CMOS circuit. Sketch the fully complementary version by augmenting the circuit with the PUN corresponding to the PDN comprising Q_5, Q_6, Q_7, and Q_8. Note that the fully complementary circuit utilizes 12 transistors. Although the circuit is more complex, it switches faster.

D 16.9 Consider another possibility for the circuit in Fig. 16.7: Relabel the R input as $\overline{S}$ and the S input as $\overline{R}$. Let $\overline{S}$ and $\overline{R}$ normally rest at V_{DD}. Let the flip-flop be storing a 0; thus $V_Q = 0$ V and $V_{\overline{Q}} = V_{DD}$. To set the flip-flop, the $\overline{S}$ terminal is lowered to 0 V and the clock ϕ is raised to V_{DD}. The relevant part of the circuit is then transistors Q_5 and Q_2. For the flip-flop to switch, the voltage at $\overline{Q}$ must be lowered to $V_{DD}/2$. What is the minimum required W/L for Q_5 in terms of $(W/L)_2$ and (μ_n/μ_p)? Assume $V_{tn} = |V_{tp}|$.

***16.10** Figure P16.10 shows a commonly used circuit of a D flip-flop that is triggered by the negative-going edge of the clock ϕ.

(a) For ϕ high, what are the values of $\overline{Q}$ and Q in terms of D? Which transistors are conducting?
(b) If D is high and ϕ goes low, which transistors conduct and what signals appear at $\overline{Q}$ and at Q? Describe the circuit operation.
(c) Repeat (b) for D low with the clock ϕ going low.
(d) Does the operation of this circuit rely on charge storage?

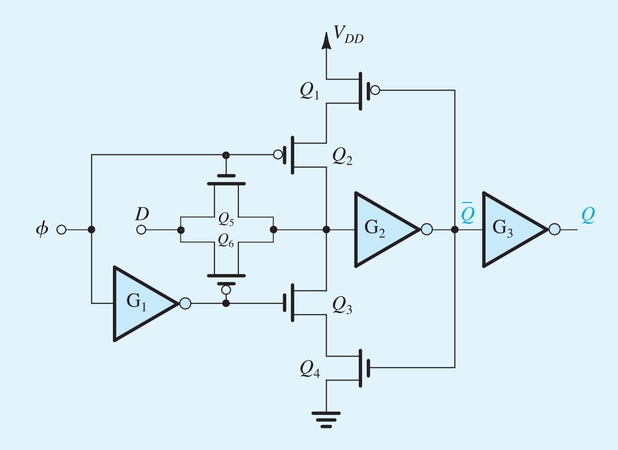


Figure P16.10

Section 16.2: Semiconductor Memories: Types and Architectures

16.11 How many cells does a 4-Gbit RAM have?

16.12 A 4-Gbit memory chip is organized as 256M words × 16 bits. How many bits does the word address need?

16.13 A particular 1-M-bit-square memory array has its peripheral circuits reorganized to allow for the readout of a 16-bit word. How many address bits will the new design need?

16.14 For the memory chip described in Problem 16.13, how many word lines must be supplied by the row decoder? How many sense amplifiers/drivers would a straightforward implementation require? If the chip power dissipation is 500 mW with a 5-V supply for continuous operation with a 20-ns cycle time, and all the power loss is dynamic, estimate the total capacitance of all logic activated in any one cycle. If we assume that 90% of this power loss occurs in array access, and that the major capacitance contributor will be the bit line itself, calculate the capacitance per bit line and per bit for this design. (Recall from Problem 16.13 that 16 bit lines are selected simultaneously.) If closer manufacturing control allows the memory array to operate at 3 V, how much larger a memory array can be designed in the same technology at about the same power level?

16.15 A 1.5-V, 1-Gbit dynamic RAM (called DRAM) by Hitachi uses a 0.16-μm process with a cell size of 0.38 × 0.76 μm² in a 19 × 38 mm² chip. What fraction of the chip is occupied by the I/O connections, peripheral circuits, and interconnect?

Section 16.3: Random-Access Memory (RAM) Cells

16.16 Repeat Exercise 16.4 for an SRAM fabricated in a 0.25-μm CMOS process for which $V_{DD} = 2.5$ V and $V_t = 0.5$ V.

16.17 Repeat Exercise 16.4 for an SRAM fabricated in a 0.13-μm CMOS process for which $V_{DD} = 1.2$ V and $V_t = 0.4$ V.

16.18 Locate on the graph of Fig. 16.14 the points A, B, and C that correspond to the following three process technologies:

(a) 0.25-μm: $V_{DD} = 2.5$ V and $V_t = 0.5$ V
(b) 0.18-μm: $V_{DD} = 1.8$ V and $V_t = 0.5$ V
(c) 0.13-μm: $V_{DD} = 1.2$ V and $V_t = 0.4$ V

In each case, impose the condition that in a read-1 operation $V_{\overline{Q}} = V_t$.

D 16.19 Find the maximum allowable W/L for the access transistors of the SRAM cell in Fig. 16.12 so that in the read operation, the voltages at Q and $\overline{Q}$ do not change by more than $|V_t|$. Assume that the SRAM is fabricated in a 0.13-μm technology for which $V_{DD} = 1.2$ V and $V_{tn} = |V_{tp}| = 0.4$ V, and $(W/L)_n = 1.5$. Find $V_{\overline{Q}}$ and I_5 that result in each of the following cases:

(i) $(W/L)_a = \frac{1}{3}$ the maximum allowed
(ii) $(W/L)_a = \frac{2}{3}$ the maximum allowed
(iii) $(W/L)_a =$ the maximum allowed

Assume $\mu_n C_{ox} = 500$ μA/V^2. Which one of the three designs results in the shortest read delay?

D 16.20 Consider a 6T SRAM cell fabricated in a 0.18-μm CMOS process for which $V_{tn} = |V_{tp}| = 0.5$ V and $V_{DD} = 1.8$ V. If during a read-1 operation it is required that $V_{\overline{Q}}$ not exceed 0.2 V, use the graph in Fig. 16.14 to determine the maximum allowable value of the ratio $(W/L)_5/(W/L)_1$. For $L_1 = L_5 = 0.18$ μm, select values for W_1 and W_5 that minimize the combined areas of Q_1 and Q_5. Assume that the minimum width allowed is 0.18 μm.

16.21 Consider the read operation of the 6T SRAM cell of Fig. 16.12 when it is storing a 0, that is, $V_Q = 0$ V, and $V_{\overline{Q}} = V_{DD}$. Assume that the bit lines are precharged to V_{DD} before the word-line voltage is raised to V_{DD}. Sketch the relevant part of the circuit and describe the operation. Show that the analysis parallels that presented in the text for the read-1 operation.

***16.22** Refer to the circuit in Fig. 16.13 and find the maximum ratio $(W/L)_5/(W/L)_1$ for $V_{\overline{Q}} \le V_t$, this time taking into account the velocity-saturation effect (Section 15.1.3, Eq. 15.11). The SRAM is fabricated in a 0.18-μm CMOS process for which $V_{DD} = 1.8$ V, $V_t = 0.5$ V, and for the n-channel devices $V_{DSsat} = 0.6$ V. Compare to the value obtained without accounting for velocity saturation. (*Hint*: Convince yourself that for this situation only Q_5 will be operating in velocity saturation.)

D *16.23 For the 6T SRAM of Fig. 16.12, fabricated in a 0.13-μm CMOS process for which $V_{DD} = 1.2$ V, $V_{t0} = 0.4$ V, $2\phi_f = 0.88$ V, and $\gamma = 0.2$ V$^{1/2}$, find the maximum ratio $(W/L)_5/(W/L)_1$ for which $V_{\overline{Q}} \le V_{t0}$ during a read-1 operation (Fig. 16.13). Then, take into account the body effect in Q_5 and compare this result to the value obtained without accounting for the body effect.

D 16.24 A 6T SRAM cell is fabricated in a 0.13-μm CMOS process for which $V_{DD} = 1.2$ V, $V_t = 0.4$ V, and $\mu_n C_{ox} = 500$ μA/V^2. The inverters utilize $(W/L)_n = 1$. Each of the bit lines has a 2-pF capacitance to ground. The sense amplifier requires a minimum of 0.2-V input for reliable and fast operation.

(a) Find the upper bound on W/L for each of the access transistors so that V_Q and $V_{\overline{Q}}$ do not change by more than V_t volts during the read operation.
(b) Find the delay time Δt encountered in the read operation if the cell design utilizes minimum-size access transistors.
(c) Find the delay time Δt if the design utilizes the maximum allowable size for the access transistors.

16.25 Consider the operation of writing a 1 into a 6T SRAM cell that is originally storing a 0. Sketch the relevant part of the circuit and explain the operation. Without doing detailed analysis, show that the analysis would lead to results identical to those obtained in the text for the write-0 operation.

D 16.26 For a 6T SRAM cell fabricated in a 0.13-μm CMOS process, find the maximum permitted value of $(W/L)_p$ in terms of $(W/L)_a$ of the access transistors. Assume $V_{DD} = 1.2$ V, $V_{tn} = |V_{tp}| = 0.4$ V, and $\mu_n = 4\mu_p$.

D 16.27 For a 6T SRAM cell fabricated in a 0.25-μm CMOS process, find the maximum permitted value of $(W/L)_p$ in terms

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

of $(W/L)_a$ of the access transistors. Assume $V_{DD} = 2.5$ V, $V_{tn} = |V_{tp}| = 0.5$ V, and $\mu_n \simeq 4\mu_p$.

16.28 Locate on the graph in Fig. 16.17 the points A, B, and C corresponding to the following three CMOS fabrication processes:

(a) 0.25-μm: $V_{DD} = 2.5$ V, $V_{tn} = |V_{tp}| = 0.5$ V
(b) 0.18-μm: $V_{DD} = 1.8$ V, $V_{tn} = |V_{tp}| = 0.5$ V
(c) 0.13-μm: $V_{DD} = 1.2$ V, $V_{tn} = |V_{tp}| = 0.4$ V

For all three, $\mu_n \simeq 4\mu_p$. In each case, V_Q is to be limited to a maximum value of V_{tn}.

D 16.29 Design a minimum-size 6T SRAM cell in a 0.13-μm process for which $V_{DD} = 1.2$ V and $V_{tn} = |V_{tp}| = 0.4$ V. All transistors are to have equal $L = 0.13$ μm. Assume that the minimum width allowed is 0.13 μm. Verify that your minimum-size cell meets the constraints in Eqs. (16.5) and (16.11).

16.30 For a particular DRAM design, the cell capacitance $C_S = 35$ fF and $V_{DD} = 1.2$ V. Each cell represents a capacitive load on the bit line of 0.8 fF. Assume a 20-fF capacitance for the sense amplifier and other circuitry attached to the bit line. What is the maximum number of cells that can be attached to a bit line while ensuring a minimum bit-line signal of 25 mV? How many bits of row addressing can be used? If the sense-amplifier gain is increased by a factor of 4, how many word-line address bits can be accommodated?

16.31 For a DRAM available for regular use 98% of the time, having a row-to-column ratio of 2 to 1, a cycle time of 10 ns, and a refresh cycle of 10 ms, estimate the total memory capacity.

16.32 In a particular dynamic memory chip, $C_S = 30$ fF, the bit-line capacitance per cell is 0.5 f F, and bit-line control circuitry involves 12 fF. For a 1-Mbit-square array, what bit-line signals result when a stored 1 is read? When a stored 0 is read? Assume that $V_{DD} = 1.2$ V.

16.33 For a DRAM cell utilizing a capacitance of 30 fF, refresh is required within 12 ms. If a signal loss on the capacitor of 0.2 V can be tolerated, what is the largest acceptable leakage current present at the cell?

Section 16.4: Sense Amplifiers and Address Decoders

D 16.34 Consider the operation of the differential sense amplifier of Fig. 16.20 following the rise of the sense control signal ϕ_s. Assume that a balanced differential signal of 0.1 V is established between the bit lines, each of which has a 1 pF capacitance. For $V_{DD} = 1.2$ V, what value of G_m of each of the inverters in the amplifier is required to cause the outputs to reach $0.1V_{DD}$ and $0.9V_{DD}$ [from initial values of $0.5V_{DD} - (0.1/2)$ and $0.5V_{DD} + (0.1/2)$ volts, respectively] in 2 ns? If for the matched inverters, $|V_t| = 0.4$ V and $k_n' = 4k_p' = 500\ \mu\text{A/V}^2$, what are the device widths required? If the input signal is 0.2 V, what does the amplifier response time become?

16.35 A particular version of the regenerative sense amplifier of Fig. 16.20 in a 0.13-μm technology uses transistors for which $|V_t| = 0.4$ V, $k_n' = 4k_p' = 500\ \mu\text{A/V}^2$, $V_{DD} = 1.2$ V, with $(W/L)_n = 0.26\ \mu\text{m}/0.13\ \mu\text{m}$ and $(W/L)_p = 1.04\ \mu\text{m}/0.13\ \mu\text{m}$. For each inverter, find the value of G_m. For a bit-line capacitance of 0.4 pF, and a delay until an output of $0.9V_{DD}$ is reached of 1 ns, find the initial difference voltage required between the two bit lines. If the time can be relaxed by 1 ns, what input signal can be handled? With the increased delay time and with the input signal at the original level, by what percentage can the bit-line capacitance, and correspondingly the bit-line length, be increased? If the delay time required for the bit-line capacitances to charge by the constant current available from the storage cell, and thus develop the difference-voltage signal needed by the sense amplifier, was 2 ns, what does it increase to when longer lines are used?

D 16.36 (a) For the sense amplifier of Fig. 16.20, show that the time required for the bit lines to reach $0.9V_{DD}$ and $0.1V_{DD}$ is given by $t_d = (C_B/G_m)\ln(0.8V_{DD}/\Delta V)$, where ΔV is the initial difference voltage between the two bit lines.
(b) If the response time of the sense amplifier is to be reduced to one-half the value of an original design, by what factor must the width of all transistors be increased?
(c) If for a particular design, $V_{DD} = 1.2$ V and $\Delta V = 0.2$ V, find the factor by which the widths of all transistors must be increased so that ΔV is reduced by a factor of 2, while keeping t_d unchanged?

D 16.37 It is required to design a sense amplifier of the type shown in Fig. 16.20 to operate with a DRAM using the dummy-cell technique illustrated in Fig. 16.22. The DRAM cell provides readout voltages of –100 mV when a 0 is stored and +40 mV when a 1 is stored. The sense amplifier is required to provide a differential output voltage of 1 V in at most 2 ns. Find the W/L ratios of the transistors in the amplifier matched inverters, assuming that the processing technology is characterized by $k_n' = 4k_p' = 300\ \mu\text{A/V}^2$, $|V_t| = 0.5$ V, and $V_{DD} = 1.8$ V. The capacitance of each half-bit line is 0.5 pF.

What will be the amplifier response time when a 0 is read? When a 1 is read?

D 16.38 It is required to design the sense amplifier of Fig. 16.24 to detect an input signal of 140 mV and provide a full output in 0.5 ns. If $C = 50$ fF and $V_{DD} = 1.2$ V, find the required current I and the power dissipation.

D 16.39 Consider the sense amplifier in Fig. 16.24 in the equilibrium condition shown in part (b) of the figure. Let $V_{DD} = 1.2$ V and $V_t = 0.4$ V.

(a) If Q_1 and Q_2 are to operate at the edge of saturation, what is the dc voltage at the drain of Q_1?
(b) If the switching voltage ΔV is to be about 140 mV, at what overdrive voltage V_{OV} should Q_1 and Q_2 be operated in equilibrium? What dc voltage should appear at the common-source terminals of Q_1 and Q_2?
(c) If the delay component Δt given by Eq. (16.18) is to be 0.5 ns, what current I is needed if $C = 55$ fF?
(d) Find the W/L required for each of Q_1 to Q_4 for $\mu_n C_{ox} = 4\mu_p C_{ox} = 500\ \mu\text{A/V}^2$.
(e) If Q_5 is to operate at the same overdrive voltage as Q_1 and Q_2, find its required W/L and the value of the reference voltage V_R,

16.40 Consider a 1024-row NOR decoder. To how many address bits does this correspond? How many output lines does the decoder have? How many input lines does the NOR array require? How many NMOS and PMOS transistors does such a design need?

16.41 For the column decoder shown in Fig. 16.26, how many column-address bits are needed in a 1-Mbit-square array? How many NMOS pass transistors are needed in the multiplexer? How many NMOS transistors are needed in the NOR decoder? How many PMOS transistors? What is the total number of NMOS and PMOS transistors needed?

16.42 Consider the use of the tree column decoder shown in Fig. 16.27 for application with a square 1-Mbit array. How many address bits are involved? How many levels of pass gates are used? How many pass transistors are there in total?

16.43 Consider a ring oscillator consisting of five inverters, each having $t_{PLH} = 3$ ns and $t_{PHL} = 2$ ns. Sketch one of the output waveforms, and specify its frequency and the percentage of the cycle during which the output is high.

16.44 A ring-of-nine oscillator is found to operate at 20 MHz. Find the propagation delay of the inverter.

D 16.45 Design the one-shot circuit of Fig. 16.29 to provide an output pulse of 10-ns width. If the inverters available have $t_P = 2.5$ ns delay, how many inverters do you need for the delay circuit?

Section 16.5: Read-Only Memory (ROM)

16.46 Give the eight words stored in the ROM of Fig. 16.30.

D 16.47 Design the bit pattern to be stored in a (16 × 4) ROM that provides the 4-bit product of two 2-bit variables. Give a circuit implementation of the ROM array using a form similar to that of Fig. 16.30.

16.48 Consider a dynamic version of the ROM in Fig. 16.30 in which the gates of the PMOS devices are connected to a precharge control signal ϕ. Let all the NMOS devices have $W/L = 3\ \mu\text{m}/1.2\ \mu\text{m}$ and all the PMOS devices have $W/L = 12\ \mu\text{m}/1.2\ \mu\text{m}$. Assume $k_n' = 3k_p' = 90\ \mu\text{A/V}^2$, $V_{tn} = -V_{tp} = 1$ V, and $V_{DD} = 5$ V.

(a) During the precharge interval, ϕ is lowered to 0 V. Estimate the time required to charge a bit line from 0 to 5 V. Use as an average charging current the current supplied by a PMOS transistor at a bit-line voltage halfway through the excursion of 0 to 5 V (i.e., 2.5 V). The bit-line capacitance is 1 pF. Note that all NMOS transistors are cut off at this time.
(b) After the precharge interval is completed and ϕ returns to V_{DD}, the row decoder raises the voltage of the selected word line. Because of the finite resistance and capacitance of the word line, the voltage rises exponentially toward V_{DD}. If the resistance of each of the polysilicon word lines is 5 kΩ and the capacitance between the word line and ground is 2 pF, what is the (10% to 90%) rise time of the word-line voltage? What is the voltage reached at the end of one time constant?
(c) If we approximate the exponential rise of the word-line voltage by a step equal to the voltage reached in one time constant, find the interval Δt required for an NMOS transistor to discharge the bit line and lower its voltage by 1 V.

Section 16.6: CMOS Image Sensors

16.49 Consider the pixel circuit in Fig. 16.34. If the capacitance C at the storage node X is 25 fF and if Q_P resets the node voltage to V_{DD}, how much electron charge is accumulated onto the capacitance when the voltage drops by 1 V? Also give the number of electrons this represents. (Recall that the magnitude of the electron charge is 1.6×10^{-19} C).

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem



PART IV

Filters and Oscillators

In Part IV we study an important class of analog circuits: filters and oscillators. Both topics have in common an application or system orientation. They provide dramatic and powerful illustration of the application of both negative and positive feedback. While the filters studied here are linear circuits, the design of oscillators makes use of both linear and nonlinear techniques.

Chapter 17 deals with the design of filters, which are important building blocks of communication and instrumentation systems. Filter design is one of the rare areas of engineering for which a complete design theory exists, starting from specification and culminating in an actual working circuit. The material presented should allow the reader to perform such a complete design process.

Chapter 17 covers most of the design methods and realization technologies currently in use for analog filters. In this respect, we should note that even though more and more signal processing is now performed digitally, analog filters are still needed. This is because most systems have to interface with the physical world, which is analog in nature. Such an interface, at a minimum, consists of an amplifier and an analog filter.

In the design of electronic systems, the need usually arises for signals of various waveforms—sinusoidal, pulse, square-wave, and so on. The generation of such signals is the subject of Chapter 18. It will be seen that some of the circuits utilized in waveform generation employ an op-amp version of the basic memory element studied in Chapter 16, the bistable multivibrator or latch.

The study of filters and oscillators relies on a thorough familiarity with basic feedback concepts, including the effect of feedback on the amplifier poles (Chapter 11), and with op-amp circuit applications (Chapter 2). As well, we assume knowledge of basic *s*-plane concepts including transfer functions (Chapter 1 and Appendix E), and poles, zeros, and Bode plots (Appendix F).

CHAPTER 17

Filters and Tuned Amplifiers

IN THIS CHAPTER YOU WILL LEARN

1. How filters are characterized by their signal-transmission properties and how they are classified into different types based on the relative location of their passband(s) and stopband(s).
2. How filters are specified and how to obtain a filter transfer function that meets the given specifications, including the use of popular special functions such as the Butterworth and the Chebyshev.
3. The various first-order and second-order filter functions and their realization using op amps and RC circuits.
4. The basic second-order LCR resonator and how it can be used to realize the various second-order filter functions.
5. The best op amp–RC circuit for realizing an inductance and how it can be used as the basis for realizing the various second-order filter functions.
6. That connecting two op-amp integrators, one inverting and one noninverting, in a feedback loop realizes a second-order resonance circuit and can be used to obtain circuit realizations of the various second-order filter functions.
7. How second-order filter functions can be realized using a single op amp and an RC circuit, and the performance limitations of these minimal realizations.
8. How the powerful concept of circuit sensitivity can be applied to assess the performance of filter circuits in the face of finite component tolerances.
9. How transconductors and capacitors can be used to design high-frequency filters in CMOS technology.
10. The basis for the most popular approach to the realization of filter functions in IC form; the switched-capacitor technique.
11. The design of tuned transistor amplifiers for radio-frequency (RF) applications.

Introduction

In this chapter, we study the design of an important building block of communications and instrumentation systems, the electronic filter. Filter design is one of the very few areas of engineering for which a complete design theory exists, starting from specification and ending

with a circuit realization. A detailed study of filter design requires an entire book, and indeed such textbooks exist. In the limited space available here, we shall concentrate on a selection of topics that provide an introduction to the subject as well as a useful arsenal of filter circuits and design methods.

The oldest technology for realizing filters makes use of inductors and capacitors, and the resulting circuits are called **passive LC filters**. Such filters work well at high frequencies; however, in low-frequency applications (dc to 100 kHz) the required inductors are large and physically bulky, and their characteristics are quite nonideal. Furthermore, such inductors are impossible to fabricate in monolithic form and are incompatible with any of the modern techniques for assembling electronic systems. Therefore, there has been considerable interest in finding filter realizations that do not require inductors. Of the various possible types of **inductorless filters**, we shall study **active-RC filters**, **transconductance-C filters**, and **switched-capacitor filters**.

Active-RC filters utilize op amps together with resistors and capacitors and are fabricated using discrete, hybrid thick-film or hybrid thin-film circuit technologies. However, for large-volume production, such technologies do not yield the economies achieved by monolithic (IC) fabrication. At the present time, there are two popular approaches for realizing fully integrated filters: the transconductance-C approach, which is particularly suited for high-frequency applications, and the switched-capacitor approach, which is used for audio-frequency applications. We shall study both methods.

The last topic studied in this chapter is the tuned amplifier commonly employed in the design of radio and TV receivers. Although tuned amplifiers are in effect bandpass filters, they are studied separately because their design is based on somewhat different techniques.

The material in this chapter requires a thorough familiarity with op-amp circuit applications. Thus the study of Chapter 2 is a prerequisite.

17.1 Filter Transmission, Types, and Specification

17.1.1 Filter Transmission

The filters we are about to study are linear circuits that can be represented by the general two-port network shown in Fig. 17.1. The filter **transfer function** $T(s)$ is the ratio of the output voltage $V_o(s)$ to the input voltage $V_i(s)$,

$$T(s) \equiv \frac{V_o(s)}{V_i(s)} \tag{17.1}$$

The filter **transmission** is found by evaluating $T(s)$ for physical frequencies, $s = j\omega$, and can be expressed in terms of its magnitude and phase as

$$T(j\omega) = |T(j\omega)|e^{j\phi(\omega)} \tag{17.2}$$

The magnitude of transmission is often expressed in decibels in terms of the **gain function**

$$G(\omega) \equiv 20 \log |T(j\omega)|, \text{dB} \tag{17.3}$$

or, alternatively, in terms of the **attenuation function**

$$A(\omega) \equiv -20 \log |T(j\omega)|, \text{dB} \tag{17.4}$$

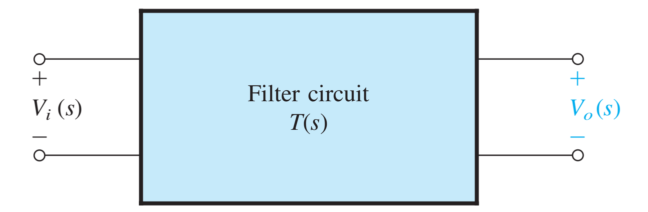


Figure 17.1 The filters studied in this chapter are linear circuits represented by the general two-port network shown. The filter transfer function $T(s) \equiv V_o(s)/V_i(s)$.

A filter shapes the frequency spectrum of the input signal, $|V_i(j\omega)|$, according to the magnitude of the transfer function $|T(j\omega)|$, thus providing an output $V_o(j\omega)$ with a spectrum

$$|V_o(j\omega)| = |T(j\omega)||V_i(j\omega)| \tag{17.5}$$

Also, the phase characteristics of the signal are modified as it passes through the filter according to the filter phase function $\phi(\omega)$.

17.1.2 Filter Types

We are specifically interested here in filters that perform a **frequency-selection** function: **passing** signals whose frequency spectrum lies within a specified range, and **stopping** signals whose frequency spectrum falls outside this range. Such a filter has ideally a frequency band (or bands) over which the magnitude of transmission is unity (the filter **passband**) and a frequency band (or bands) over which the transmission is zero (the filter **stopband**). Figure 17.2 depicts the ideal transmission characteristics of the four major filter types: **low-pass** (LP) in Fig. 17.2(a), **high-pass** (HP) in Fig. 17.2(b), **bandpass** (BP) in Fig. 17.2(c), and **bandstop** (BS) or **band-reject** in Fig. 17.2(d). These idealized characteristics, by virtue of their vertical edges, are known as **brick-wall** responses.

17.1.3 Filter Specification

The filter-design process begins with the filter user specifying the transmission characteristics required of the filter. Such a specification cannot be of the form shown in Fig. 17.2 because physical circuits cannot realize these idealized characteristics. Figure 17.3 shows realistic specifications for the transmission characteristics of a low-pass filter. Observe that since a physical circuit cannot provide constant transmission at all passband frequencies, the specifications allow for deviation of the passband transmission from the ideal 0 dB, but place an upper bound, A_{max} (dB), on this deviation. Depending on the application, A_{max} typically ranges from 0.05 dB to 3 dB. Also, since a physical circuit cannot provide zero transmission at all stopband frequencies, the specifications in Fig. 17.3 allow for some transmission over the stopband. However, the specifications require the stopband signals to be attenuated by at least A_{min} (dB) relative to the passband signals. Depending on the filter application, A_{min} can range from 20 dB to 100 dB.

Since the transmission of a physical circuit cannot change abruptly at the edge of the passband, the specifications of Fig. 17.3 provide for a band of frequencies over which the attenuation increases from near 0 dB to A_{min}. This **transition band** extends from the passband edge ω_p to the stopband edge ω_s. The ratio ω_s/ω_p is usually used as a measure of the sharpness of the low-pass filter response and is called the **selectivity factor**. Finally, observe that for convenience the passband transmission is specified to be 0 dB. The final filter, however, can be given a passband gain, if desired, without changing its selectivity characteristics.

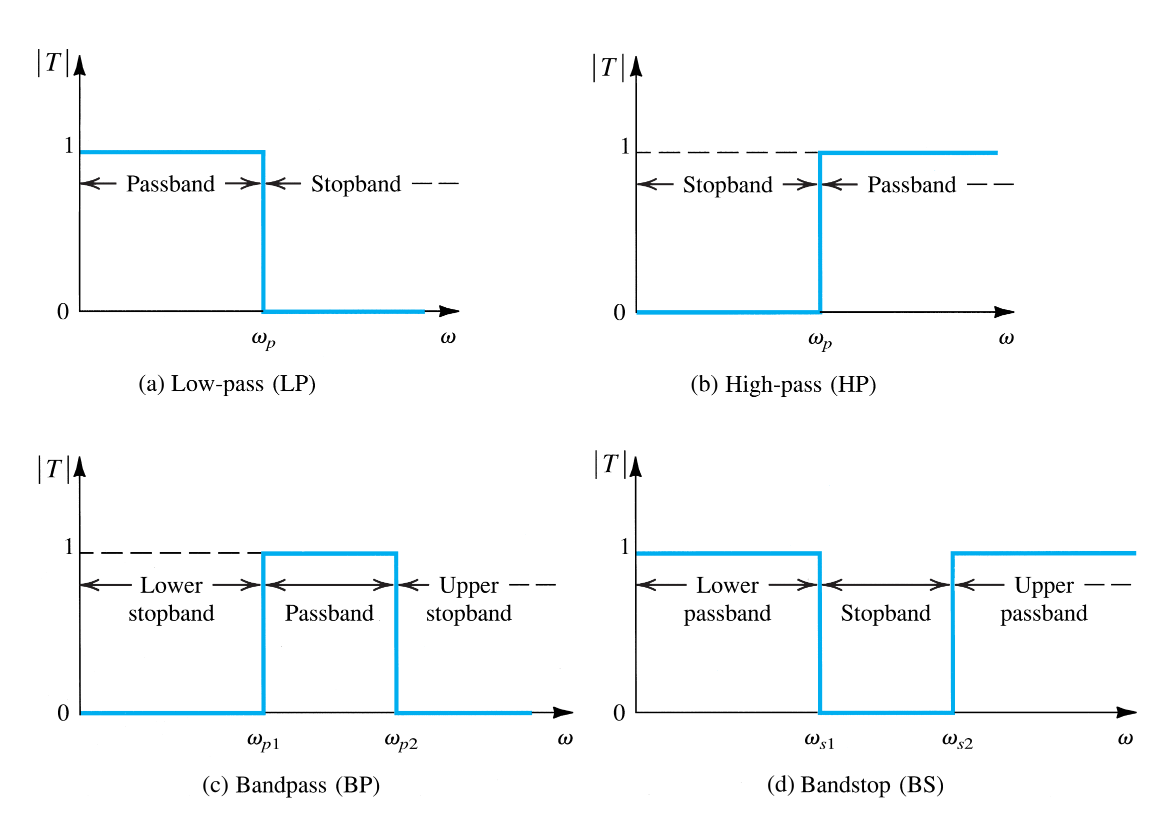


Figure 17.2 Ideal transmission characteristics of the four major filter types: **(a)** low-pass (LP), **(b)** high-pass (HP), **(c)** bandpass (BP), and **(d)** bandstop (BS).

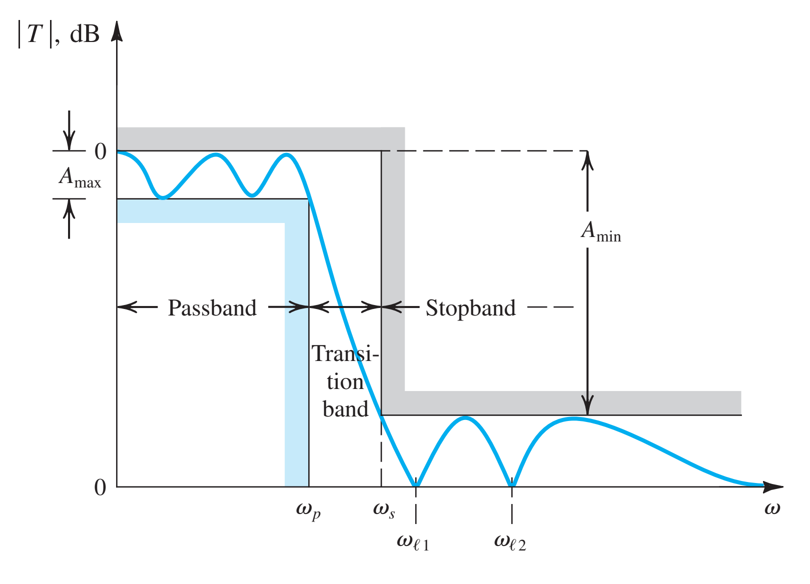


Figure 17.3 Specification of the transmission characteristics of a low-pass filter. The magnitude response of a filter that just meets specifications is also shown.

A BRIEF HISTORY OF ANALOG FILTERS:

Driven by the needs of the emerging telephone system in the early part of the twentieth century, the earliest analog filters utilized inductors (coils) and capacitors and were designed using ad hoc methods. It was not until the 1940s that a design theory for passive LC filters was developed. This filter synthesis method, however, required extensive computation and had to await the emergence of the digital computer and its widespread use in the 1950s and the 1960s to become adopted, further developed, and utilized.

The invention of the transistor and the IC resulted in new ways of assembling electronic circuits, which were, however, incompatible with the bulky and heavy inductors used in audio-frequency LC filters. The search for inductorless filters thus began and was greatly influenced by the emergence of the IC op amp in the late 1960s and its availability at low cost. Extensive research in the 1960s and 1970s resulted in a large repertoire of op amp–RC filters, the best of which are studied in this chapter.

No sooner had op amp–RC filters become a mature and reliable technology than the need arose for analog filters that could be fully integrated. Multiple approaches were proposed in the late 1970s, but the switched-capacitor circuit has become the technology of choice for low-frequency applications, mostly because of its compatibility with CMOS. For higher-frequency applications approaching the gigahertz range, needed for the burgeoning mobile communications market, two approaches have gained momentum: transconductance-C and active-LC. The latter uses actual physical inductors (in the nanohenry range) fabricated on the IC chip.

To summarize, the transmission of a low-pass filter is specified by four parameters:

1. The passband edge ω_p
2. The maximum allowed variation in passband transmission A_{max}
3. The stopband edge ω_s
4. The minimum required stopband attenuation A_{min}

The more tightly one specifies a filter—that is, lower A_{max}, higher A_{min}, and/or a selectivity ratio ω_s/ω_p closer to unity—the closer the response of the resulting filter will be to the ideal. However, the resulting filter circuit will be of higher order and thus more complex and expensive.

In addition to the magnitude of transmission, the phase response of the filter is of interest in some applications. The filter-design problem, however, is considerably complicated when both magnitude and phase are to be specified.

Once the filter specifications have been decided upon, the next step in the design is to find a transfer function whose magnitude meets the specification. To meet specification, the magnitude-response curve must lie in the unshaded area in Fig. 17.3. The curve shown in the figure is for a filter that *just* meets specifications. Observe that for this particular filter, the magnitude response *ripples* throughout the passband, and the ripple peaks are all equal. Since the peak ripple is equal to A_{max} it is usual to refer to A_{max} as the **passband ripple** and to ω_p as the **ripple bandwidth**. The particular filter response shows ripples also in the stopband, again with the ripple peaks all equal and of such a value that the minimum stopband attenuation achieved is equal to the specified value, A_{min}. Thus this particular response is said to be **equiripple** in both the passband and the stopband.

The process of obtaining a transfer function that meets given specifications is known as **filter approximation**. Filter approximation is usually performed using computer programs (Snelgrove, 1982; Ouslis and Sedra, 1995) or filter-design tables (Zverev, 1967). In simpler

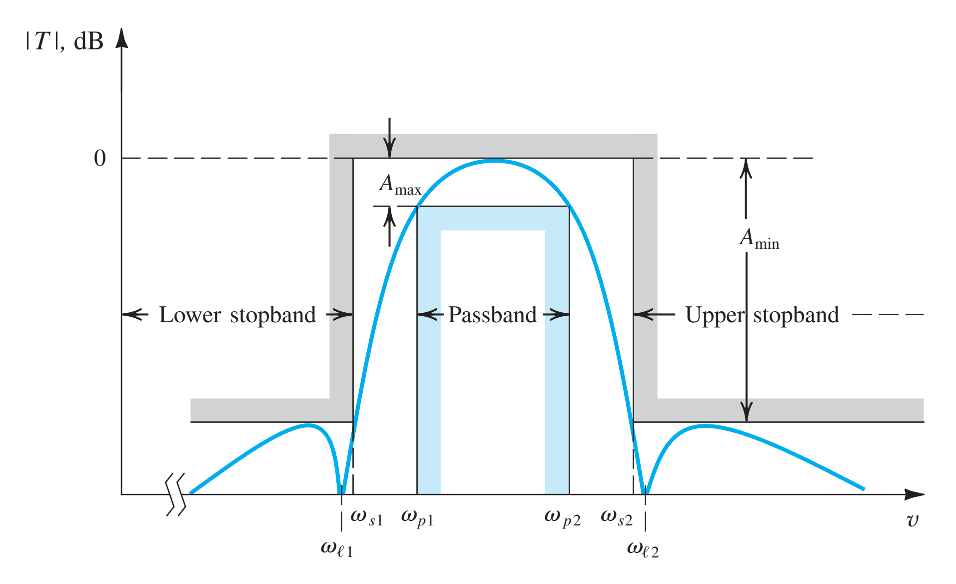


Figure 17.4 Transmission specifications for a bandpass filter. The magnitude response of a filter that just meets specifications is also shown. Note that this particular filter has a monotonically decreasing transmission in the passband on both sides of the peak frequency.

cases, filter approximation can be performed using closed-form expressions, as will be seen in Section 17.3.

Finally, Fig. 17.4 shows transmission specifications for a bandpass filter and the response of a filter that meets these specifications. For this example we have chosen an approximation function that does not ripple in the passband; rather, the transmission decreases monotonically on both sides of the center frequency, attaining the maximum allowable deviation at the two edges of the passband.

EXERCISES

17.1 Find approximate values of attenuation (in dB) corresponding to filter transmissions of 1, 0.99, 0.9, 0.8, 0.7, 0.5, 0.1, 0.
Ans. 0, 0.1, 1, 2, 3, 6, 20, ∞ (dB)

17.2 If the magnitude of passband transmission is to remain constant to within ±5%, and if the stopband transmission is to be no greater than 1% of the passband transmission, find A_{max} and A_{min}.
Ans. 0.9 dB; 40 dB

17.2 The Filter Transfer Function

The filter transfer function $T(s)$ can be written as the ratio of two polynomials as

$$T(s) = \frac{a_M s^M + a_{M-1}s^{M-1} + \cdots + a_0}{s^N + b_{N-1}s^{N-1} + \cdots + b_0} \tag{17.6}$$

The degree of the denominator, N, is the **filter order**. For the filter circuit to be stable, the degree of the numerator must be less than or equal to that of the denominator: $M \le N$. The numerator and denominator coefficients, $a_0, a_1, \ldots, a_M$ and $b_0, b_1, \ldots, b_{N-1}$, are real numbers. The polynomials in the numerator and denominator can be factored, and $T(s)$ can be expressed in the form

$$T(s) = \frac{a_M(s-z_1)(s-z_2)\cdots(s-z_M)}{(s-p_1)(s-p_2)\cdots(s-p_N)} \tag{17.7}$$

The numerator roots, $z_1, z_2, \ldots, z_M$, are the **transfer function zeros**, or **transmission zeros**; and the denominator roots, $p_1, p_2, \ldots, p_N$, are the **transfer function poles**, or the **natural modes**.[1] Each transmission zero or pole can be either a real or a complex number. Complex zeros and poles, however, must occur in conjugate pairs. Thus, if $-1+j2$ happens to be a zero, then $-1-j2$ also must be a zero.

Since in the filter stopband the transmission is required to be zero or small, the filter transmission zeros are usually placed on the $j\omega$ axis at stopband frequencies. This indeed is the case for the filter whose transmission function is sketched in Fig. 17.3. This particular filter can be seen to have zero transmission (infinite attenuation) at two stopband frequencies: ω_{l1} and ω_{l2}. The filter then must have transmission zeros at $s = +j\omega_{l1}$ and $s = +j\omega_{l2}$. However, since complex zeros occur in conjugate pairs, there must also be transmission zeros at $s = -j\omega_{l1}$ and $s = -j\omega_{l2}$. Thus the numerator polynomial of this filter will have the factors $(s+j\omega_{l1})(s-j\omega_{l1})(s+j\omega_{l2})(s-j\omega_{l2})$, which can be written as $\left(s^2+\omega_{l1}^2\right)\left(s^2+\omega_{l2}^2\right)$. For $s = j\omega$ (physical frequencies) the numerator becomes $\left(-\omega^2+\omega_{l1}^2\right)\left(-\omega^2+\omega_{l2}^2\right)$, which indeed is zero at $\omega = \omega_{l1}$ and $\omega = \omega_{l2}$.

Continuing with the example in Fig. 17.3, we observe that the transmission decreases toward zero as ω approaches ∞. Thus the filter must have one or more transmission zeros at $s = \infty$. In general, the number of transmission zeros at $s = \infty$ is the difference between the degree of the numerator polynomial, M, and the degree of the denominator polynomial, N, of the transfer function in Eq. (17.6). This is because as s approaches ∞, $T(s)$ approaches a_M/s^{N-M} and thus is said to have $N-M$ zeros at $s = \infty$.

For a filter circuit to be stable, all its poles must lie in the left half of the s plane, and thus p_1, $p_2, \ldots, p_N$ must all have negative real parts. Figure 17.5 shows typical pole and zero locations for the low-pass filter whose transmission function is depicted in Fig. 17.3. We have assumed that this filter is of fifth order ($N = 5$). It has two pairs of complex-conjugate poles and one real-axis pole, for a total of five poles. All the poles lie in the vicinity of the passband, which is what gives the filter its high transmission at passband frequencies. The five transmission zeros are at $s = \pm j\omega_{l1}, s = \pm j\omega_{l2}$, and $s = \infty$. Thus, the transfer function for this filter is of the form

$$T(s) = \frac{a_4\left(s^2+\omega_{l1}^2\right)\left(s^2+\omega_{l2}^2\right)}{s^5+b_4s^4+b_3s^3+b_2s^2+b_1s+b_0} \tag{17.8}$$

As another example, consider the bandpass filter whose magnitude response is shown in Fig. 17.4. This filter has transmission zeros at $s = \pm j\omega_{l1}$ and $s = \pm j\omega_{l2}$. It also has one or more zeros at $s = 0$ and one or more zeros at $s = \infty$ (because the transmission decreases toward 0 as ω approaches 0 and ∞). Assuming that only one zero exists at each of $s = 0$ and $s = \infty$,

[1]Throughout this chapter, we use the names *poles* and *natural modes* interchangeably.

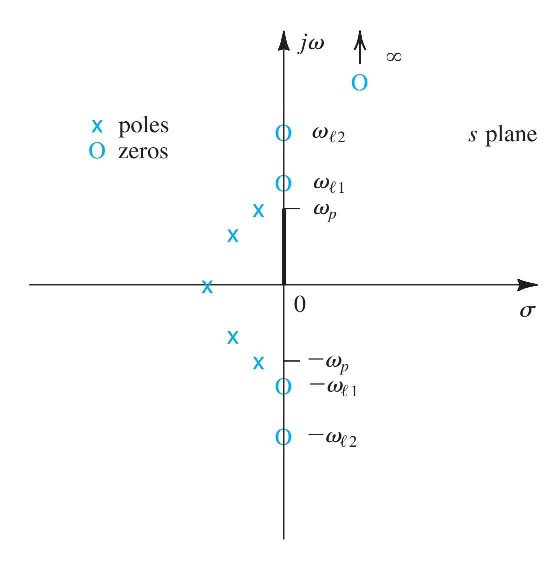


Figure 17.5 Pole–zero pattern for the low-pass filter whose transmission is sketched in Fig. 17.3. This is a fifth-order filter ($N = 5$).

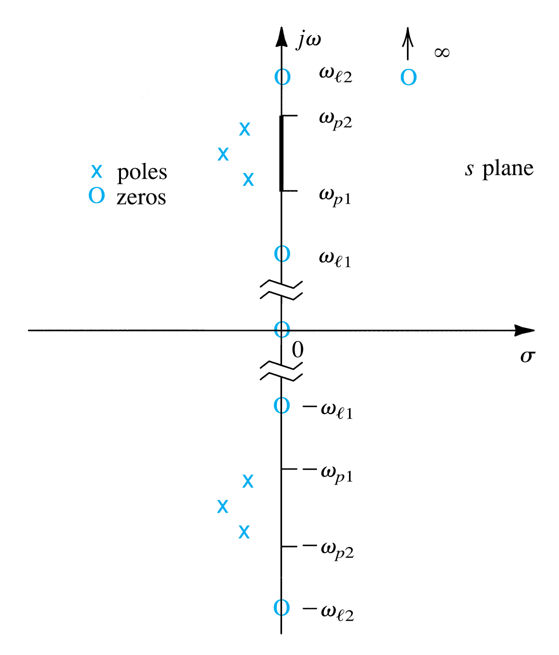


Figure 17.6 Pole–zero pattern for the band-pass filter whose transmission function is shown in Fig. 17.4. This is a sixth-order filter ($N = 6$).

the filter must be of sixth order, and its transfer function takes the form

$$T(s) = \frac{a_5 s\left(s^2 + \omega_{l1}^2\right)\left(s^2 + \omega_{l2}^2\right)}{s^6 + b_5 s^5 + \ldots + b_0} \tag{17.9}$$

A typical pole–zero plot for such a filter is shown in Fig. 17.6.

As a third and final example, consider the low-pass filter whose transmission function is depicted in Fig. 17.7(a). We observe that in this case there are no finite values of ω at which the transmission is zero. Thus it is possible that all the transmission zeros of this filter are at

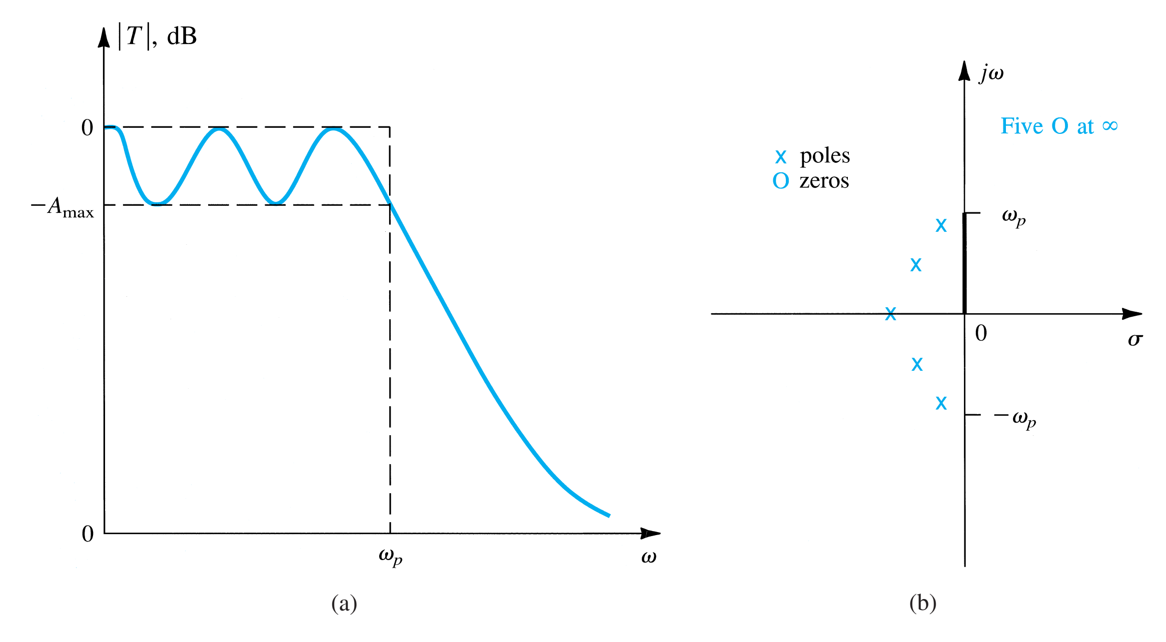


Figure 17.7 **(a)** Transmission characteristics of a fifth-order low-pass filter having all transmission zeros at infinity. **(b)** Pole–zero pattern for the filter in **(a)**.

$s = \infty$. If this is the case, the filter transfer function takes the form

$$T(s) = \frac{a_0}{s^N + b_{N-1}s^{N-1} + \ldots + b_0} \tag{17.10}$$

Such a filter is known as an **all-pole filter**. Typical pole–zero locations for a fifth-order all-pole low-pass filter are shown in Fig. 17.7(b).

Almost all the filters studied in this chapter have all their transmission zeros on the $j\omega$ axis, in the filter stopband(s), including[2] $\omega = 0$ and $\omega = \infty$. Also, to obtain high selectivity, all the natural modes will be complex conjugate (except for the case of odd-order filters, where one natural mode must be on the real axis). Finally we note that the more selective the required filter response is, the higher its order must be, and the closer its natural modes are to the $j\omega$ axis.

EXERCISES

17.3 A second-order filter has its poles at $s = -(1/2) \pm j\left(\sqrt{3}/2\right)$. The transmission is zero at $\omega = 2$ rad/s and is unity at dc ($\omega = 0$). Find the transfer function.

Ans. $T(s) = \frac{1}{4}\frac{s^2+4}{s^2+s+1}$

[2]Obviously, a low-pass filter should *not* have a transmission zero at $\omega = 0$, and, similarly, a high-pass filter should not have a transmission zero at $\omega = \infty$.

17.4 A fourth-order filter has zero transmission at $\omega = 0$, $\omega = 2$ rad/s, and $\omega = \infty$. The natural modes are $-0.1 \pm j0.8$ and $-0.1 \pm j1.2$. Find $T(s)$.

Ans. $T(s) = \dfrac{a_3 s\left(s^2 + 4\right)}{(s^2 + 0.2s + 0.65)(s^2 + 0.2s + 1.45)}$

17.5 Find the transfer function $T(s)$ of a third-order all-pole low-pass filter whose poles are at a radial distance of 1 rad/s from the origin and whose complex poles are at 30° angles from the $j\omega$ axis. The dc gain is unity. Show that $|T(j\omega)| = 1/\sqrt{1+\omega^6}$. Find ω_{3dB} and the attenuation at $\omega = 3$ rad/s.
Ans. $T(s) = 1/(s+1)\ (s^2 + s + 1)$; 1 rad/s; 28.6 dB

17.3 Butterworth and Chebyshev Filters

In this section, we present two functions that are frequently used in approximating the transmission characteristics of low-pass filters; that is, in obtaining a transfer function $T(s)$ whose magnitude $|T(j\omega)|$ meets given low-pass filter specifications. Closed-form expressions are available for the parameters of these functions, and thus one can use them in filter design without the need for computers or filter-design tables. Their utility, however, is limited to relatively simple applications.

Although in this section we discuss the design of low-pass filters only, the approximation functions presented can be applied to the design of other filter types through the use of frequency transformations (see Sedra and Brackett, 1978).

17.3.1 The Butterworth Filter

Figure 17.8 shows a sketch of the magnitude response of a Butterworth[3] filter. This filter exhibits a monotonically decreasing transmission with all the transmission zeros at $\omega = \infty$, making it an all-pole filter. The magnitude function for an Nth-order Butterworth filter with a passband edge ω_p is given by

$$|T(j\omega)| = \frac{1}{\sqrt{1 + \epsilon^2 \left(\dfrac{\omega}{\omega_p}\right)^{2N}}} \tag{17.11}$$

At $\omega = \omega_p$,

$$\left|T\left(j\omega_p\right)\right| = \frac{1}{\sqrt{1+\epsilon^2}} \tag{17.12}$$

Thus, the parameter ϵ determines the maximum variation in passband transmission, A_{max}, according to

$$A_{max} = 20 \log \sqrt{1+\epsilon^2} \tag{17.13}$$

[3] The Butterworth filter approximation is named after S. Butterworth, a British engineer who in 1930 was among the first to employ it.

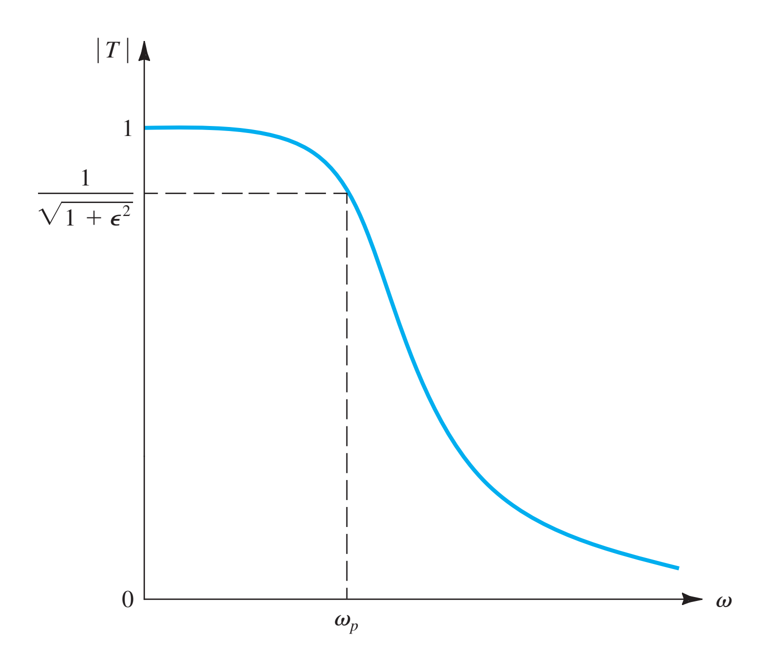


Figure 17.8 The magnitude response of a Butterworth filter.

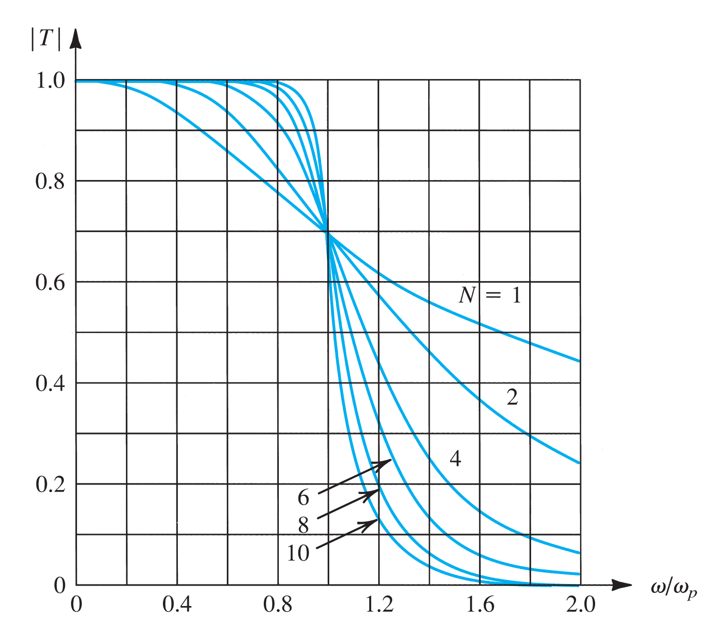


Figure 17.9 Magnitude response for Butterworth filters of various order with $\epsilon = 1$. Note that as the order increases, the response approaches the ideal brick-wall type of transmission.

Conversely, given A_{max}, the value of ϵ can be determined from

$$\epsilon = \sqrt{10^{A_{max}/10} - 1} \tag{17.14}$$

Observe that in the Butterworth response the maximum deviation in passband transmission (from the ideal value of unity) occurs at the passband edge only. It can be shown that the

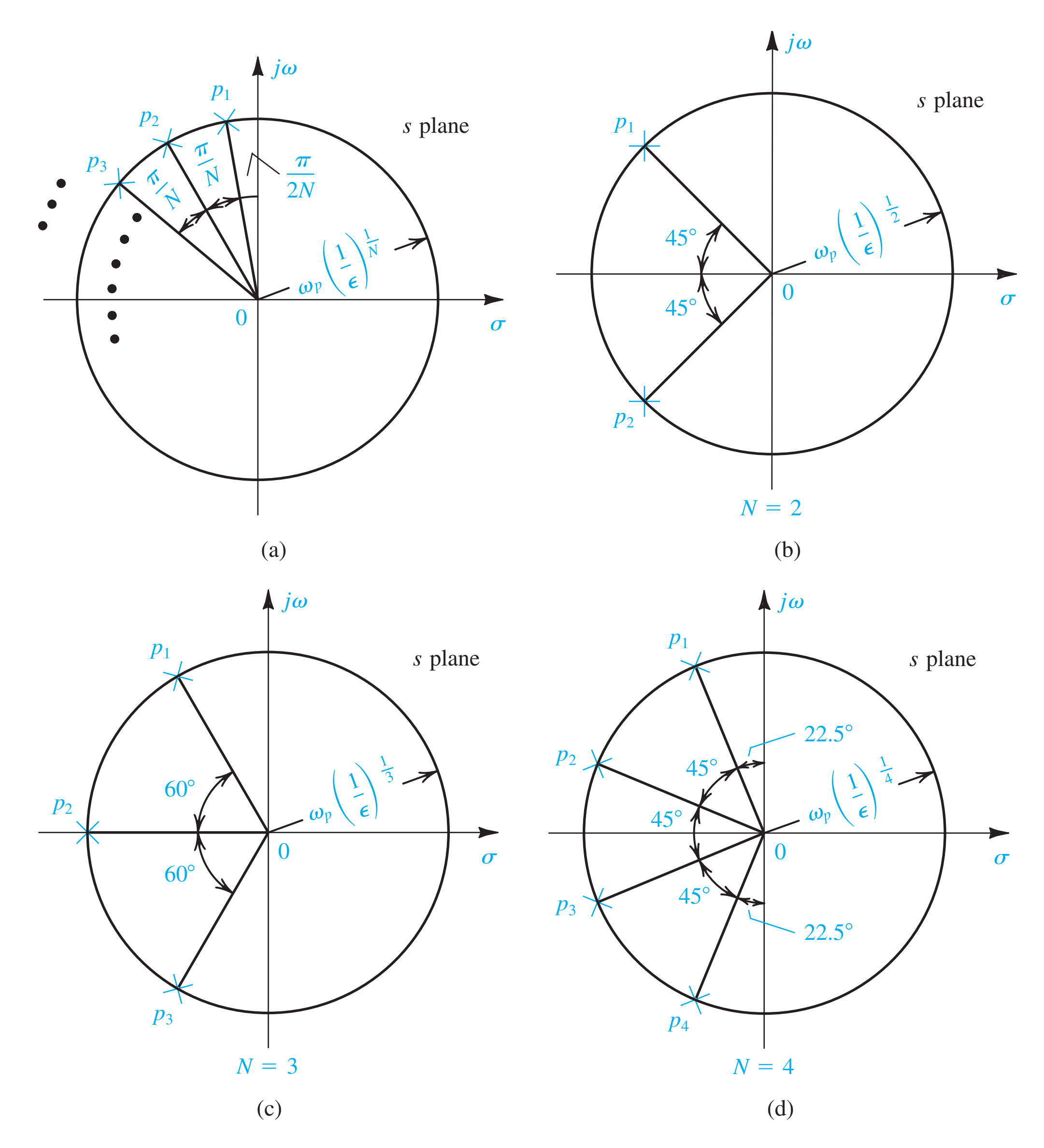


Figure 17.10 Graphical construction for determining the poles of a Butterworth filter of order N. All the poles lie in the left half of the s plane on a circle of radius $\omega_0 = \omega_p(1/\epsilon)^{1/N}$, where ϵ is the passband deviation parameter $\left(\epsilon = \sqrt{10^{A_{\max}/10} - 1}\right)$: **(a)** the general case; **(b)** $N = 2$; **(c)** $N = 3$; **(d)** $N = 4$.

first $2N - 1$ derivatives of $|T|$ relative to ω are zero at $\omega = 0$ (see Van Valkenburg, 1980). This property makes the Butterworth response very flat near $\omega = 0$ and results in the name **maximally flat** response. The degree of passband flatness increases as the order N is increased, as can be seen from Fig. 17.9. This figure indicates also that, as should be expected, as the order N is increased, the filter response approaches the ideal brick-wall type of response.

At the edge of the stopband, $\omega = \omega_s$, the attenuation of the Butterworth filter can be obtained by substituting $\omega = \omega_s$ in Eq. (17.11). The result is given by

$$A(\omega_s) = -20\log\left[1\Big/\sqrt{1+\epsilon^2(\omega_s/\omega_p)^{2N}}\right] \tag{17.15}$$

$$= 10\log\left[1+\epsilon^2(\omega_s/\omega_p)^{2N}\right]$$

This equation can be used to determine the filter order required, which is the lowest integer value of N that yields $A(\omega_s) \geq A_{\min}$.

The natural modes of an Nth-order Butterworth filter can be determined from the graphical construction shown in Fig. 17.10(a). Observe that the natural modes lie on a circle of radius $\omega_p(1/\epsilon)^{1/N}$ and are spaced by equal angles of π/N, with the first mode at an angle $\pi/2N$ from the $+j\omega$ axis. Since the natural modes all have equal radial distance from the origin, they all have the same frequency $\omega_0 = \omega_p(1/\epsilon)^{1/N}$. See Fig. 17.10(b), (c), and (d) for the natural modes of Butterworth filters of order $N = 2, 3$, and 4, respectively. Once the N natural modes p_1, $p_2, \ldots, p_N$ have been found, the transfer function can be written as

$$T(s) = \frac{K\omega_0^N}{(s-p_1)(s-p_2)\cdots(s-p_N)} \tag{17.16}$$

where K is a constant equal to the required dc gain of the filter.

To summarize, to find a Butterworth transfer function that meets transmission specifications of the form in Fig. 17.3 we perform the following procedure:

1. Determine ϵ from Eq. (17.14).
2. Use Eq. (17.15) to determine the required filter order as the lowest integer value of N that results in $A(\omega_s) \geq A_{\min}$.
3. Use Fig. 17.10(a) to determine the N natural modes.
4. Use Eq. (17.16) to determine $T(s)$

Example 17.1

Find the Butterworth transfer function that meets the following low-pass filter specifications: $f_p = 10$ kHz, $A_{\max} = 1$ dB, $f_s = 15$ kHz, $A_{\min} = 25$ dB, dc gain $= 1$.

Solution

Substituting $A_{\max} = 1$ dB into Eq. (17.14) yields $\epsilon = 0.5088$. Equation (17.15) is then used to determine the filter order by trying various values for N. We find that $N = 8$ yields $A(\omega_s) = 22.3$ dB and $N = 9$ gives 25.8 dB. We thus select $N = 9$.

Figure 17.11 shows the graphical construction for determining the poles. The poles all have the same frequency $\omega_0 = \omega_p(1/\epsilon)^{1/N} = 2\pi \times 10 \times 10^3(1/0.5088)^{1/9} = 6.773 \times 10^4$ rad/s. The first pole p_1 is given by

$$p_1 = \omega_0(-\cos\ 80° + j\ \sin\ 80°) = \omega_0(-0.1736 + j0.9848)$$

Example 17.1 *continued*

Combining p_1 with its complex conjugate p_9 yields the factor $\left(s^2+s0.3472\,\omega_0+\omega_0^2\right)$ in the denominator of the transfer function. The same can be done for the other complex poles, and the complete transfer function is obtained using Eq. (17.16),

$$T(s)=\frac{\omega_0^9}{(s+\omega_0)(s^2+s1.8794\omega_0+\omega_0^2)(s^2+s1.5321\omega_0+\omega_0^2)}\times\frac{1}{(s^2+s\omega_0+\omega_0^2)(s^2+s0.3472\omega_0+\omega_0^2)} \tag{17.17}$$

17.3.2 The Chebyshev Filter

Figure 17.12 shows representative transmission functions for Chebyshev[4] filters of even and odd orders. The Chebyshev filter exhibits an equiripple response in the passband and a monotonically decreasing transmission in the stopband. While the odd-order filter has $|T(0)|=1$, the even-order filter exhibits its maximum magnitude deviation at $\omega=0$. In both cases the total number of passband maxima and minima equals the order of the filter, N.

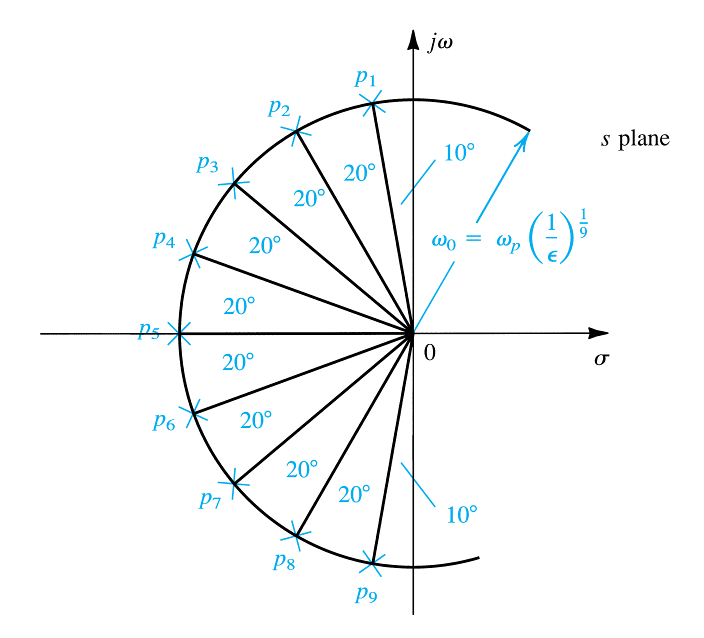


Figure 17.11 Poles of the ninth-order Butterworth filter of Example 17.1.

[4]Named after the Russian mathematician P. L. Chebyshev, who in 1899 used these functions in studying the construction of steam engines.

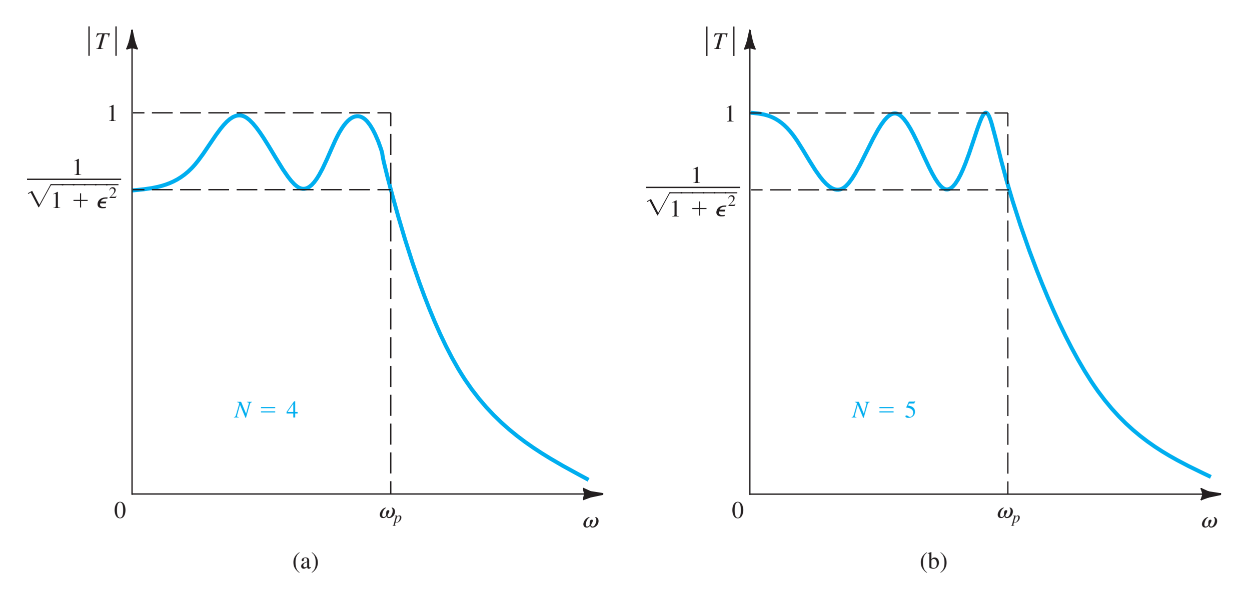


Figure 17.12 Sketches of the transmission characteristics of representative **(a)** even-order and **(b)** odd-order Chebyshev filters.

All the transmission zeros of the Chebyshev filter are at $\omega = \infty$, making it an all-pole filter.

The magnitude of the transfer function of an Nth-order Chebyshev filter with a passband edge (ripple bandwidth) ω_p is given by

$$|T(j\omega)| = \frac{1}{\sqrt{1+\epsilon^2\cos^2\left[N\cos^{-1}(\omega/\omega_p)\right]}} \quad \text{for} \quad \omega \le \omega_p \tag{17.18}$$

and

$$|T(j\omega)| = \frac{1}{\sqrt{1+\epsilon^2\cosh^2\left[N\cosh^{-1}(\omega/\omega_p)\right]}} \quad \text{for} \quad \omega \ge \omega_p \tag{17.19}$$

At the passband edge, $\omega = \omega_p$, the magnitude function is given by

$$|T(j\omega_p)| = \frac{1}{\sqrt{1+\epsilon^2}}$$

Thus, the parameter ϵ determines the passband ripple according to

$$A_{\max} = 10\log(1+\epsilon^2) \tag{17.20}$$

Conversely, given $A_{\max}$, the value of ϵ is determined from

$$\epsilon = \sqrt{10^{A_{\max}/10}-1} \tag{17.21}$$

The attenuation achieved by the Chebyshev filter at the stopband edge ($\omega = \omega_s$) is found using Eq. (17.19) as

$$A(\omega_s) = 10\log\left[1+\epsilon^2\cosh^2\left(N\cosh^{-1}(\omega_s/\omega_p)\right)\right] \tag{17.22}$$

With the aid of a calculator, this equation can be used to determine the order N required to obtain a specified A_{min} by finding the lowest integer value of N that yields $A(\omega_s) \geq A_{min}$. As in the case of the Butterworth filter, increasing the order N of the Chebyshev filter causes its magnitude function to approach the ideal brick-wall low-pass response.

The poles of the Chebyshev filter are given by

$$p_k = -\omega_p \sin\left(\frac{2k-1}{N}\frac{\pi}{2}\right)\sinh\left(\frac{1}{N}\sinh^{-1}\frac{1}{\epsilon}\right) + j\omega_p \cos\left(\frac{2k-1}{N}\frac{\pi}{2}\right)\cosh\left(\frac{1}{N}\sinh^{-1}\frac{1}{\epsilon}\right) \qquad k = 1, 2, \ldots, N \tag{17.23}$$

Finally, the transfer function of the Chebyshev filter can be written as

$$T(s) = \frac{K\omega_p^N}{\epsilon 2^{N-1}(s-p_1)(s-p_2)\cdots(s-p_N)} \tag{17.24}$$

where K is the dc gain that the filter is required to have.

To summarize, given low-pass transmission specifications of the type shown in Fig. 17.3, the transfer function of a Chebyshev filter that meets these specifications can be found as follows:

1. Determine ϵ from Eq. (17.21).
2. Use Eq. (17.22) to determine the order required.
3. Determine the poles using Eq. (17.23).
4. Determine the transfer function using Eq. (17.24).

The Chebyshev filter provides a more efficient approximation than the Butterworth filter. Thus, for the same order and the same A_{max}, the Chebyshev filter provides greater stopband attenuation than the Butterworth filter. Alternatively, to meet identical specifications, one requires a lower order for the Chebyshev than for the Butterworth filter. This point will be illustrated by the following example.

Example 17.2

Find the Chebyshev transfer function that meets the same low-pass filter specifications given in Example 17.1: namely, $f_p = 10$ kHz, $A_{max} = 1$ dB, $f_s = 15$ kHz, $A_{min} = 25$ dB, dc gain = 1.

Solution

Substituting $A_{max} = 1$ dB into Eq. (17.21) yields $\epsilon = 0.5088$. By trying various values for N in Eq. (17.22) we find that $N = 4$ yields $A(\omega_s) = 21.6$ dB and $N = 5$ provides 29.9 dB. We thus select $N = 5$. Recall that we required a ninth-order Butterworth filter to meet the same specifications in Example 17.1.

The poles are obtained by substituting in Eq. (17.23) as

$$p_1, p_5 = \omega_p(-0.0895 \pm j0.9901)$$

$$p_2, p_4 = \omega_p(-0.2342 \pm j0.6119)$$

$$p_5 = \omega_p(-0.2895)$$

The transfer function is obtained by substituting these values in Eq. (17.24) as

$$T(s) = \frac{\omega_p^5}{8.1408(s + 0.2895\,\omega_p)(s^2 + s0.4684\,\omega_p + 0.4293\,\omega_p^2)} \times \frac{1}{s^2 + s\,0.1789\,\omega_p + 0.9883\,\omega_p^2} \quad (17.25)$$

where $\omega_p = 2\pi \times 10^4$ rad/s.

EXERCISES

D17.6 Determine the order N of a Butterworth filter for which $A_{max} = 1$ dB, $\omega_s/\omega_p = 1.5$, and $A_{min} = 30$ dB. What is the actual value of minimum stopband attenuation realized? If A_{min} is to be exactly 30 dB, to what value can A_{max} be reduced?
Ans. $N = 11$; $A_{min} = 32.87$ dB; 0.54 dB

17.7 Find the natural modes and the transfer function of a Butterworth filter with $\omega_p = 1$ rad/s, $A_{max} = 3$ dB ($\epsilon \simeq 1$), and $N = 3$.
Ans. $-0.5 \pm j\sqrt{3}/2$ and -1; $T(s) = 1/(s+1)\left(s^2 + s + 1\right)$

17.8 Observe that Eq. (17.18) can be used to find the frequencies in the passband at which $|T|$ is at its peaks and at its valleys. (The peaks are reached when the $\cos^2[\ ]$ term is zero, and the valleys correspond to the $\cos^2[\ ]$ term equal to unity.) Find these frequencies for a fifth-order filter.
Ans. Peaks at $\omega = 0$, $0.59\omega_p$, and $0.95\omega_p$; the valleys at $\omega = 0.31\omega_p$ and $0.81\omega_p$

D17.9 Find the attenuation provided at $\omega = 2\omega_p$ by a seventh-order Chebyshev filter with a 0.5-dB passband ripple. If the passband ripple is allowed to increase to 1 dB, by how much does the stopband attenuation increase?
Ans. 64.9 dB; 3.3 dB

D17.10 It is required to design a low-pass filter having $f_p = 1$ kHz, $A_{max} = 1$ dB, $f_s = 1.5$ kHz, $A_{min} = 50$ dB. (a) Find the required order of a Chebyshev filter. What is the excess stopband attenuation obtained? (b) Repeat for a Butterworth filter.
Ans. (a) $N = 8$, 5 dB; (b) $N = 16$, 0.5 dB

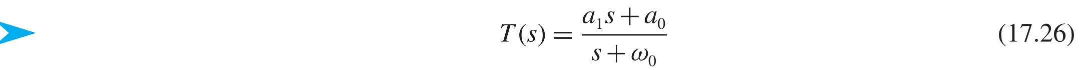
17.4 First-Order and Second-Order Filter Functions

In this section, we shall study the simplest filter transfer functions, those of first and second order. These functions are useful in their own right in the design of simple filters. First- and second-order filters can also be cascaded to realize a high-order filter. Cascade design is in fact one of the most popular methods for the design of active filters (those utilizing op amps and

RC circuits). Because the filter poles occur in complex-conjugate pairs, a high-order transfer function $T(s)$ is factored into the product of second-order functions. If $T(s)$ is odd, there will also be a first-order function in the factorization. Each of the second-order functions [and the first-order function when $T(s)$ is odd] is then realized using one of the op amp–RC circuits that will be studied in this chapter, and the resulting blocks are placed in cascade. If the output of each block is taken at the output terminal of an op amp where the impedance level is low (ideally zero), cascading does not change the transfer functions of the individual blocks. Thus the overall transfer function of the cascade is simply the product of the transfer functions of the individual blocks, which is the original $T(s)$.

17.4.1 First-Order Filters

The general first-order transfer function is given by



$$T(s) = \frac{a_1 s + a_0}{s + \omega_0} \tag{17.26}$$

This **bilinear transfer function** characterizes a first-order filter with a pole at $s = -\omega_0$, a transmission zero at $s = -a_0/a_1$, and a high-frequency gain that approaches a_1. The numerator coefficients, a_0 and a_1, determine the type of filter (e.g., low pass, high pass, etc.). Some special cases together with passive (RC) and active (op amp–RC) realizations are shown in Fig. 17.13.[5] Note that the active realizations provide considerably more versatility than their passive counterparts; in many cases the gain can be set to a desired value, and some transfer function parameters can be adjusted without affecting others. The output impedance of the active circuit is also very low, making cascading easily possible. The op amp, however, limits the high-frequency operation of the active circuits.

An important special case of the first-order filter function is the **all-pass filter** shown in Fig. 17.14. Here, the transmission zero and the pole are symmetrically located relative to the $j\omega$ axis. (They are said to display mirror-image symmetry with respect to the $j\omega$ axis.) Observe that although the transmission of the all-pass filter is (ideally) constant at all frequencies, its phase shows frequency selectivity. All-pass filters are used as phase shifters and in systems that require phase shaping (e.g., in the design of circuits called *delay equalizers,* which cause the overall time delay of a transmission system to be constant with frequency).

EXERCISES

D17.11 Using $R_1 = 10$ kΩ, design the op amp–RC circuit of Fig. 17.13(b) to realize a high-pass filter with a corner frequency of 10^4 rad/s and a high-frequency gain of 10.
Ans. $R_2 = 100$ kΩ; $C = 0.01$ μF

D17.12 Design the op amp–RC circuit of Fig. 17.14 to realize an all-pass filter with a 90° phase shift at 10^3 rad/s. Select suitable component values.
Ans. Possible choices: $R = R_1 = R_2 = 10$ kΩ; $C = 0.1$ μF

[5] This figure contains a wealth of information about first-order filters; we recommend that the reader study it carefully.

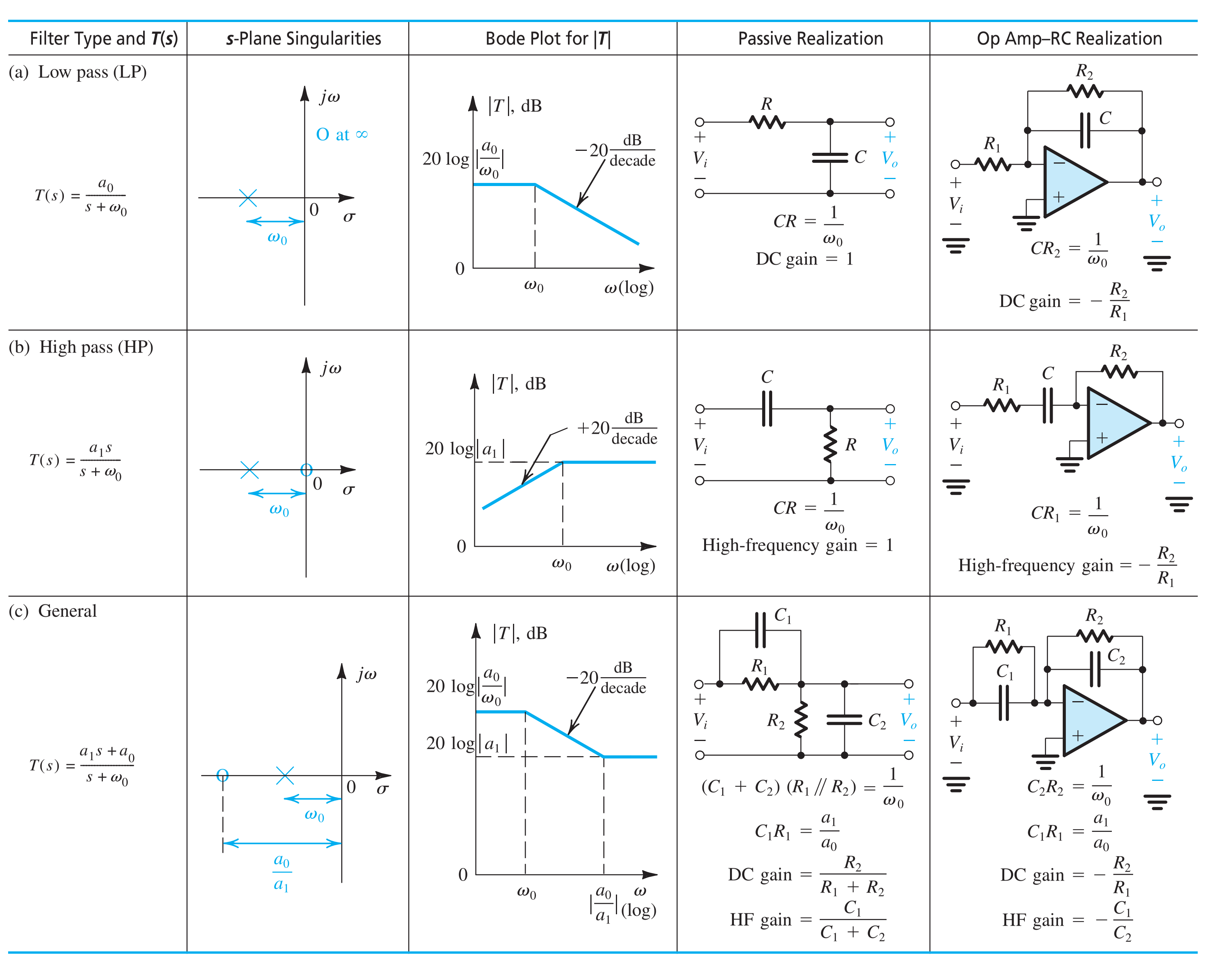


Figure 17.13 First-order filters.

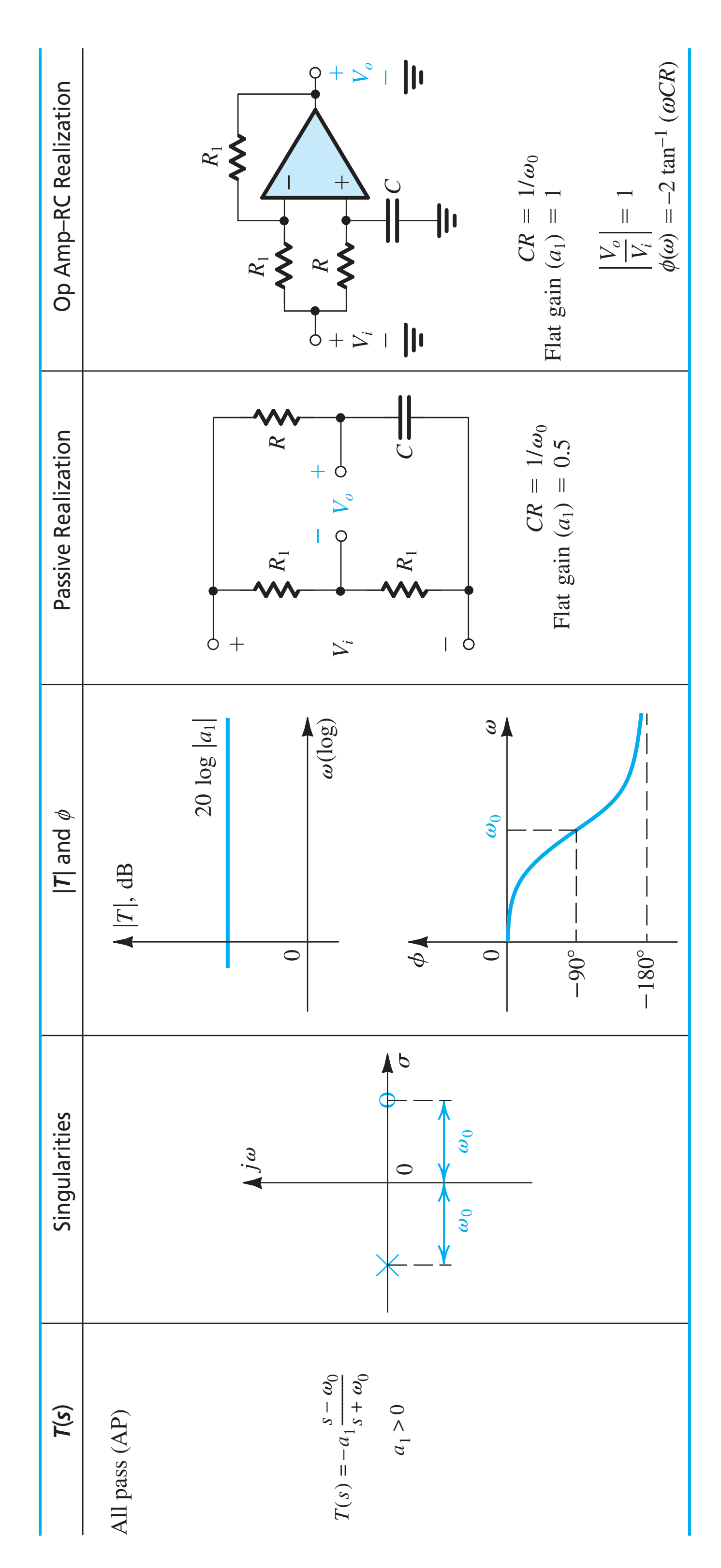


Figure 17.14 First-order all-pass filter.

17.4.2 Second-Order Filter Functions

The general second-order (or **biquadratic**) filter transfer function is usually expressed in the standard form

$$T(s) = \frac{a_2 s^2 + a_1 s + a_0}{s^2 + (\omega_0/Q)s + \omega_0^2} \tag{17.27}$$

where ω_0 and Q determine the natural modes (poles) according to

$$p_1, p_2 = -\frac{\omega_0}{2Q} \pm j\omega_0\sqrt{1 - (1/4Q^2)} \tag{17.28}$$

We are usually interested in the case of complex-conjugate natural modes, obtained for $Q > 0.5$. Figure 17.15 shows the location of the pair of complex-conjugate poles in the s plane. Observe that the radial distance of the natural modes (from the origin) is equal to ω_0, which is known as the **pole frequency**. The parameter Q determines the distance of the poles from the $j\omega$ axis: the higher the value of Q, the closer the poles are to the $j\omega$ axis, and the more selective the filter response becomes. An infinite value for Q locates the poles on the $j\omega$ axis and can yield sustained oscillations in the circuit realization. A negative value of Q implies that the poles are in the right half of the s plane, which certainly produces oscillations. The parameter Q is called the **pole quality factor**, or simply **pole Q**.

The transmission zeros of the second-order filter are determined by the numerator coefficients, a_0, a_1, and a_2. It follows that the numerator coefficients determine the type of second-order filter function (i.e., LP, HP, etc.). Seven special cases of interest are illustrated in Fig. 17.16. For each case we give the transfer function, the s-plane locations of the transfer function singularities, and the magnitude response. Circuit realizations for the various second-order filter functions will be given in subsequent sections.

All seven special second-order filters have a pair of complex-conjugate natural modes characterized by a frequency ω_0 and a quality factor Q.

In the low-pass (LP) case, shown in Fig. 17.16(a), the two transmission zeros are at $s = \infty$. The magnitude response can exhibit a peak with the details indicated. It can be shown that the peak occurs only for $Q > 1/\sqrt{2}$. The response obtained for $Q = 1/\sqrt{2}$ is the Butterworth, or maximally flat, response.

The high-pass (HP) function shown in Fig. 17.16(b) has both transmission zeros at $s = 0$ (dc). The magnitude response shows a peak for $Q > 1/\sqrt{2}$, with the details of the response as indicated. Observe the duality between the LP and HP responses.

Next consider the bandpass (BP) filter function shown in Fig. 17.16(c). Here, one transmission zero is at $s = 0$ (dc), and the other is at $s = \infty$. The magnitude response peaks at $\omega = \omega_0$. Thus the **center frequency** of the bandpass filter is equal to the pole frequency ω_0. The

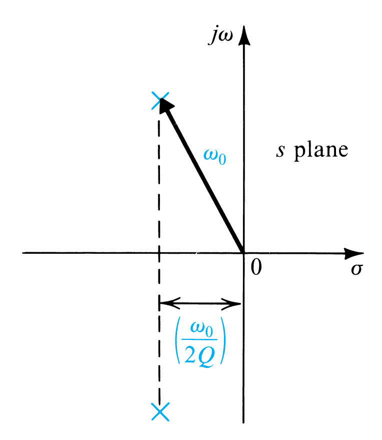


Figure 17.15 Definition of the parameters ω_0 and Q of a pair of complex-conjugate poles.

selectivity of the second-order bandpass filter is usually measured by its *3-dB bandwidth*. This is the difference between the two frequencies ω_1 and ω_2 at which the magnitude response is 3 dB below its maximum value (at ω_0). It can be shown that

$$\omega_1, \omega_2 = \omega_0\sqrt{1+(1/4Q^2)} \pm \frac{\omega_0}{2Q} \tag{17.29}$$

Thus,

$$BW \equiv \omega_2 - \omega_1 = \omega_0/Q \tag{17.30}$$

Observe that as Q increases, the bandwidth decreases and the bandpass filter becomes more selective.

If the transmission zeros are located on the $j\omega$ axis, at the complex-conjugate locations $\pm j\omega_n$, then the magnitude response exhibits zero transmission at $\omega = \omega_n$. Thus a **notch** in the magnitude response occurs at $\omega = \omega_n$, and ω_n is known as the **notch frequency**. Three cases of the second-order notch filter are possible: the regular notch, obtained when $\omega_n = \omega_0$ [Fig. 17.16(d)]; the low-pass notch, obtained when $\omega_n > \omega_0$ [Fig. 17.16(e)]; and the high-pass notch, obtained when $\omega_n < \omega_0$ [Fig. 17.16(f)]. The reader is urged to verify the response details given in these figures (a rather tedious task, though!). Observe that in all notch cases, the transmission at dc and at $s = \infty$ is finite. This is so because there are no transmission zeros at either $s = 0$ or $s = \infty$.

The last special case of interest is the all-pass (AP) filter whose characteristics are illustrated in Fig. 17.16(g). Here the two transmission zeros are in the right half of the s plane, at the mirror-image locations of the poles. (This is the case for all-pass functions of any order.) The magnitude response of the all-pass function is constant over all frequencies; the **flat gain**, as it is called, is in our case equal to $|a_2|$. The frequency selectivity of the all-pass function is in its phase response.

EXERCISES

17.13 For a maximally flat second-order low-pass filter ($Q = 1/\sqrt{2}$), show that at $\omega = \omega_0$ the magnitude response is 3 dB below the value at dc.

17.14 Give the transfer function of a second-order bandpass filter with a center frequency of 10^5 rad/s, a center-frequency gain of 10, and a 3-dB bandwidth of 10^3 rad/s.

Ans. $T(s) = \dfrac{10^4 s}{s^2 + 10^3 s + 10^{10}}$

17.15 (a) For the second-order notch function with $\omega_n = \omega_0$, show that for the attenuation to be greater than A dB over a frequency band BW_a, the value of Q is given by

$$Q \le \frac{\omega_0}{BW_a\sqrt{10^{A/10}-1}}$$

(*Hint:* First, show that any two frequencies, ω_1 and ω_2, at which $|T|$ is the same are related by $\omega_1\omega_2 = \omega_0^2$.)

(b) Use the result of (a) to show that the 3-dB bandwidth is ω_0/Q, as indicated in Fig. 17.16(d).

17.16 Consider a low-pass notch with $\omega_0 = 1$ rad/s, $Q = 10$, $\omega_n = 1.2$ rad/s, and a dc gain of unity. Find the frequency and magnitude of the transmission peak. Also find the high-frequency transmission.

Ans. 0.986 rad/s; 3.17; 0.69

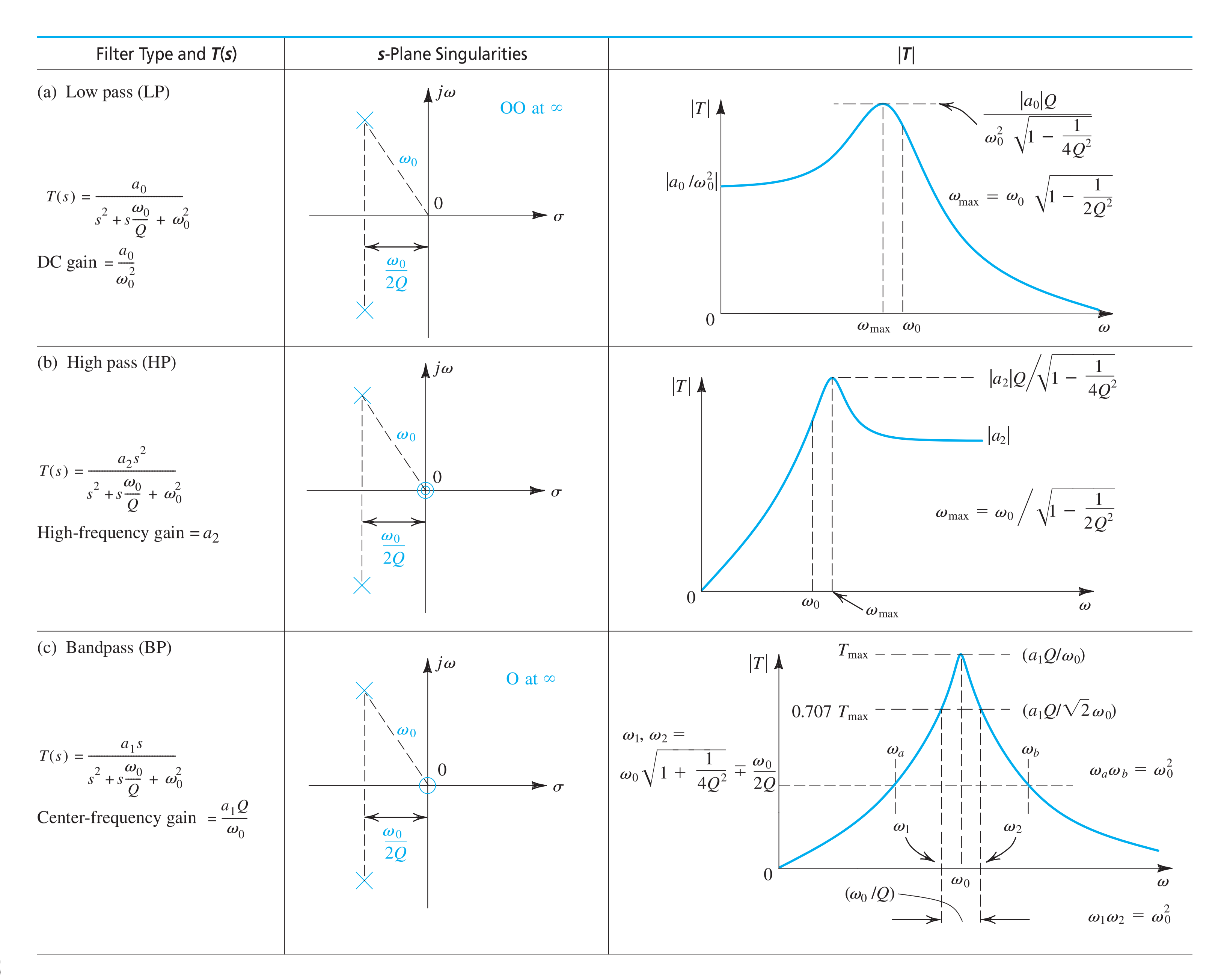


Figure 17.16 Second-order filtering functions.

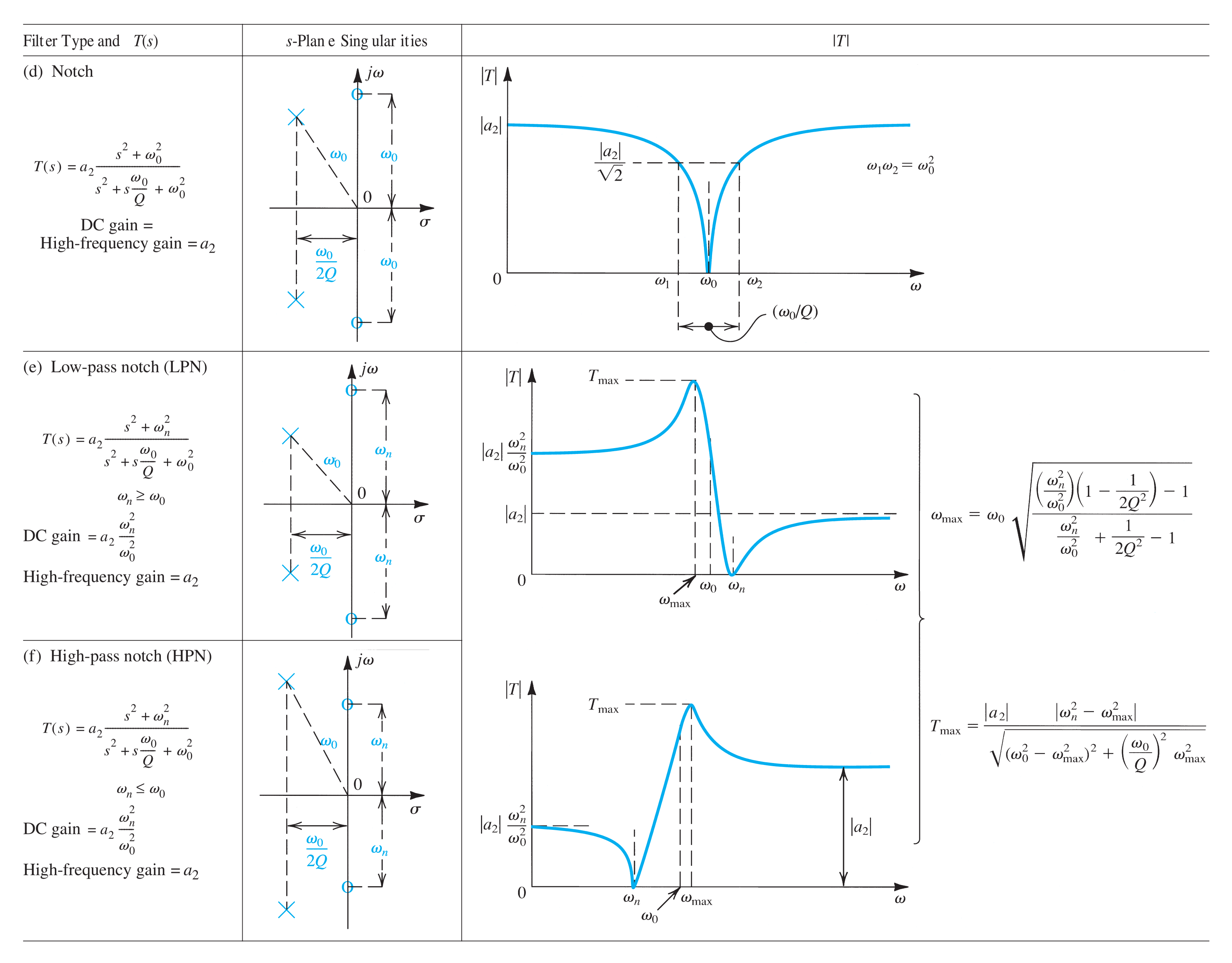


Figure 17.16 *continued*

(g) All pass (AP)

$$T(s) = a_2 \frac{s^2 - s\frac{\omega_0}{Q} + \omega_0^2}{s^2 + s\frac{\omega_0}{Q} + \omega_0^2}$$

Flat gain $= a_2$

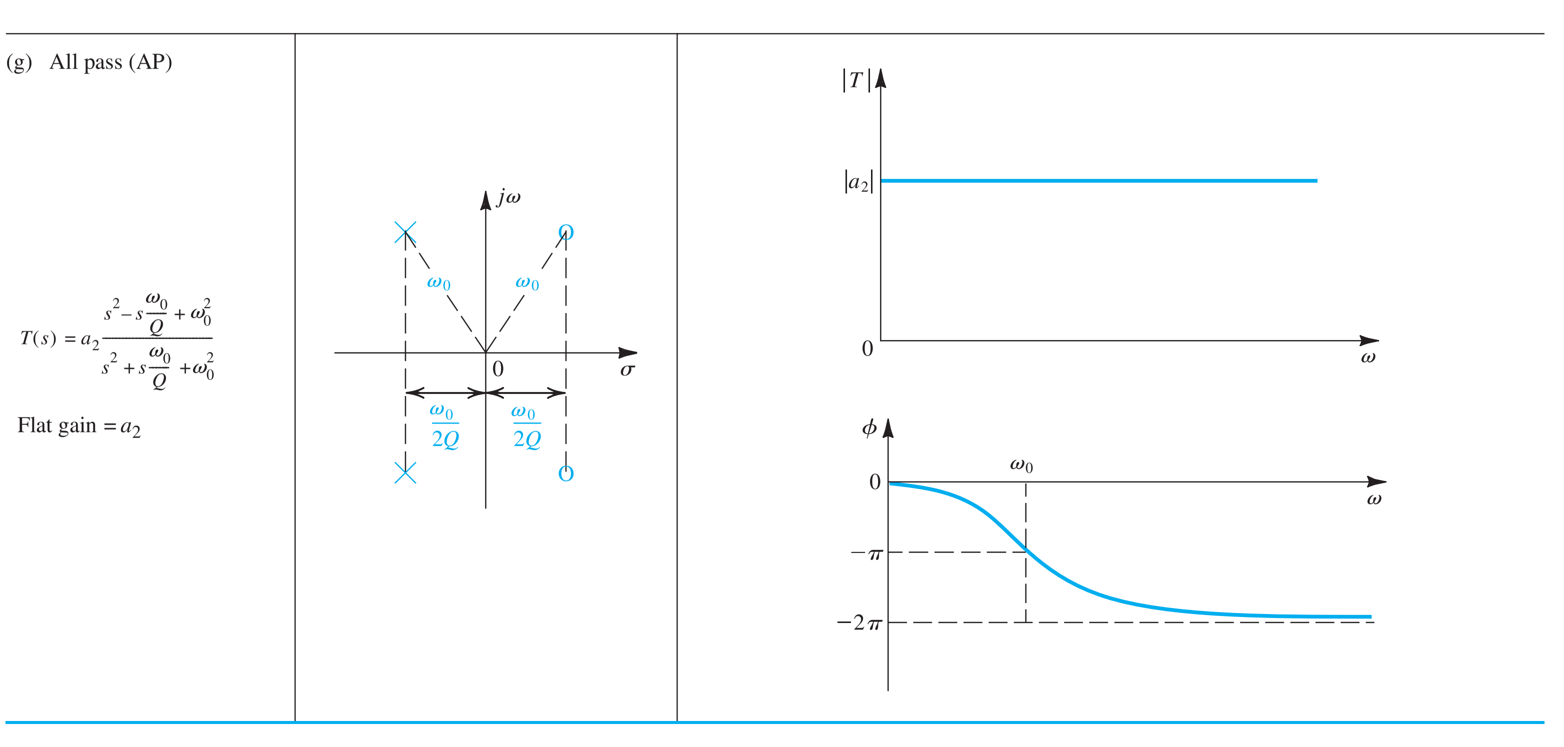


Figure 17.16 *continued*

17.5 The Second-Order LCR Resonator

In this section we shall study the second-order LCR resonator shown in Fig. 17.17(a). The use of this resonator to derive circuit realizations for the various second-order filter functions will be demonstrated. It will be shown in the next section that replacing the inductor L by a simulated inductance obtained using an op amp–RC circuit results in an op amp–RC resonator. The latter forms the basis of an important class of active-RC filters to be studied in Section 17.6.

17.5.1 The Resonator Natural Modes

The natural modes or poles of the parallel resonance circuit of Fig. 17.17(a) can be determined by applying *an excitation that does not change the natural structure of the circuit*. Two possible ways of exciting the circuit are shown in Fig. 17.17(b) and (c). In Fig. 17.17(b) the resonator is excited with a current source I connected in parallel. Since, as far as the natural response of a circuit is concerned, an independent ideal current source is equivalent to an open circuit, the excitation of Fig. 17.17(b) does not alter the natural structure of the resonator. Thus the circuit in Fig. 17.17(b) can be used to determine the natural modes of the resonator by simply finding the poles of any response function. We can for instance take the voltage V_o across the resonator as the response and thus obtain the response function $V_o/I = Z$, where Z is the impedance of the parallel resonance circuit. However, because of the parallel structure of the circuit it is more convenient to work in terms of the admittance Y; thus,

$$\frac{V_o}{I} = \frac{1}{Y} = \frac{1}{(1/sL) + sC + (1/R)} \tag{17.31}$$

$$= \frac{s/C}{s^2 + s(1/CR) + (1/LC)}$$

Equating the denominator to the standard form $\left[s^2 + s(\omega_0/Q) + \omega_0^2\right]$ leads to

$$\omega_0^2 = 1/LC \tag{17.32}$$

and

$$\omega_0/Q = 1/CR \tag{17.33}$$

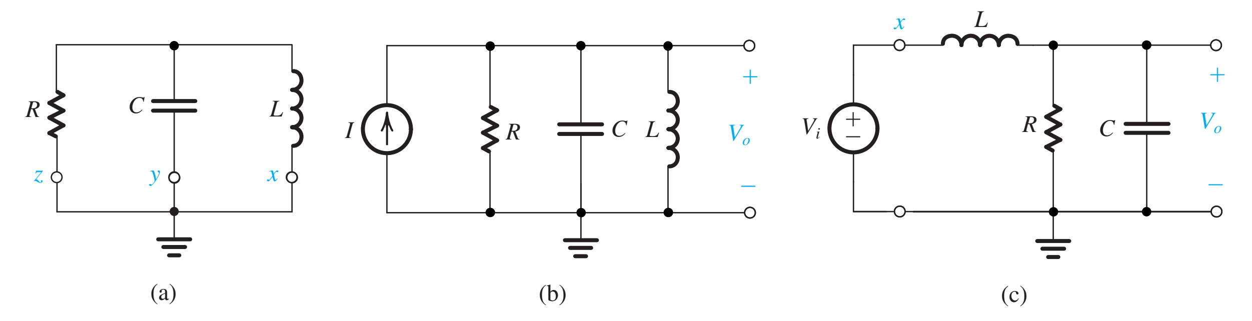


Figure 17.17 **(a)** The second-order parallel LCR resonator. **(b, c)** Two ways of exciting the resonator of **(a)** without changing its *natural structure*; resonator poles are those poles of V_o/I and V_o/V_i.

Thus,

$$\omega_0 = 1/\sqrt{LC} \tag{17.34}$$

$$Q = \omega_0 CR \tag{17.35}$$

These expressions should be familiar to the reader from studies of parallel resonance circuits in introductory courses on circuit theory.

An alternative way of exciting the parallel LCR resonator for the purpose of determining its natural modes is shown in Fig. 17.17(c). Here, node x of inductor L has been disconnected from ground and connected to an ideal voltage source V_i. Now, since as far as the natural response of a circuit is concerned, an ideal independent voltage source is equivalent to a short circuit, the excitation of Fig. 17.17(c) does not alter the natural structure of the resonator. Thus we can use the circuit in Fig. 17.17(c) to determine the natural modes of the resonator. These are the poles of any response function. For instance, we can select V_o as the response variable and find the transfer function V_o/V_i. The reader can easily verify that this will lead to the natural modes determined earlier.

In a design problem, we will be given ω_0 and Q and will be asked to determine L, C, and R. Equations (17.34) and (17.35) are two equations in the three unknowns. The one available degree of freedom can be utilized to set the impedance level of the circuit to a value that results in practical component values.

17.5.2 Realization of Transmission Zeros

Having selected the component values of the LCR resonator to realize a given pair of complex-conjugate natural modes, we now consider the use of the resonator to realize a desired filter type (e.g., LP, HP, etc.). Specifically, we wish to find out where to inject the input voltage signal V_i so that the transfer function V_o/V_i is the desired one. Toward that end, note that in the resonator circuit in Fig. 17.17(a), any of the nodes labeled x, y, or z can be disconnected from ground and connected to V_i without altering the circuit's natural modes. When this is done, the circuit takes the form of a voltage divider, as shown in Fig. 17.18(a). Thus the transfer function realized is

$$T(s) = \frac{V_o(s)}{V_i(s)} = \frac{Z_2(s)}{Z_1(s) + Z_2(s)} \tag{17.36}$$

We observe that *the transmission zeros are the values of s at which $Z_2(s)$ is zero, provided $Z_1(s)$ is not simultaneously zero, and the values of s at which $Z_1(s)$ is infinite, provided $Z_2(s)$ is not simultaneously infinite*. This statement makes physical sense: The output will be zero either when $Z_2(s)$ behaves as a short circuit or when $Z_1(s)$ behaves as an open circuit. If there is a value of s at which both Z_1 and Z_2 are zero, then V_o/V_i will be finite and no transmission zero is obtained. Similarly, if there is a value of s at which both Z_1 and Z_2 are infinite, then V_o/V_i will be finite and no transmission zero is realized.

17.5.3 Realization of the Low-Pass Function

Using the scheme just outlined, we see that to realize a low-pass function, node x is disconnected from ground and connected to V_i, as shown in Fig. 17.18(b). The transmission zeros of this circuit will be at the value of s for which the series impedance becomes

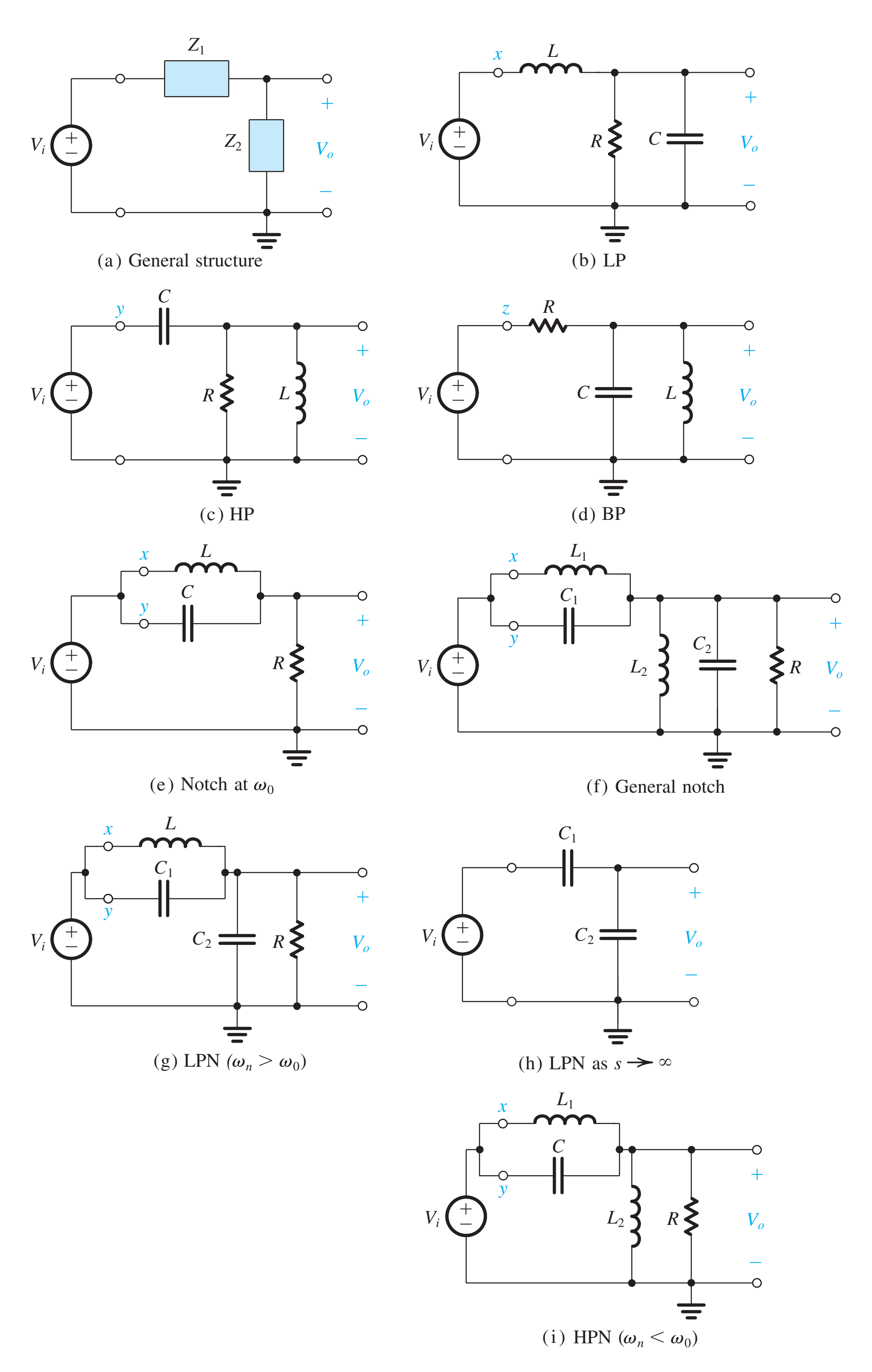


Figure 17.18 Realization of various second-order filter functions using the LCR resonator of Fig. 17.17(b): **(a)** general structure, **(b)** LP, **(c)** HP, **(d)** BP, **(e)** notch at ω_0, **(f)** general notch, **(g)** LPN $(\omega_n \geq \omega_0)$, **(h)** LPN as $s \to \infty$, **(i)** HPN $(\omega_n < \omega_0)$.

infinite (sL becomes infinite at $s = \infty$) and the value of s at which the shunt impedance becomes zero ($1/[sC + (1/R)]$ becomes zero at $s = \infty$). Thus this circuit has two transmission zeros at $s = \infty$, as a second-order LP is supposed to. The transfer function can be written either by inspection or by using the voltage divider rule. Following the latter approach, we obtain

$$T(s) \equiv \frac{V_o}{V_i} = \frac{Z_2}{Z_1 + Z_2} = \frac{Y_1}{Y_1 + Y_2} = \frac{1/sL}{(1/sL) + sC + (1/R)}$$

$$= \frac{1/LC}{s^2 + s(1/CR) + (1/LC)} \tag{17.37}$$

17.5.4 Realization of the High-Pass Function

To realize the second-order high-pass function, node y is disconnected from ground and connected to V_i, as shown in Fig. 17.18(c). Here the series capacitor introduces a transmission zero at $s = 0$ (dc), and the shunt inductor introduces another transmission zero at $s = 0$ (dc). Thus, by inspection, the transfer function may be written as

$$T(s) \equiv \frac{V_o}{V_i} = \frac{a_2 s^2}{s^2 + s(\omega_0/Q) + \omega_0^2} \tag{17.38}$$

where ω_0 and Q are the natural mode parameters given by Eqs. (17.34) and (17.35) and a_2 is the high-frequency transmission. The value of a_2 can be determined from the circuit by observing that as s approaches ∞, the capacitor approaches a short circuit and V_o approaches V_i, resulting in $a_2 = 1$.

17.5.5 Realization of the Bandpass Function

The bandpass function is realized by disconnecting node z from ground and connecting it to V_i, as shown in Fig. 17.18(d). Here the series impedance is resistive and thus does not introduce any transmission zeros. These are obtained as follows: One zero at $s = 0$ is realized by the shunt inductor, and one zero at $s = \infty$ is realized by the shunt capacitor. At the center frequency ω_0, the parallel LC-tuned circuit exhibits an infinite impedance, and thus no current flows in the circuit. It follows that at $\omega = \omega_0$, $V_o = V_i$. In other words, the center-frequency gain of the bandpass filter is unity. Its transfer function can be obtained as follows:

$$T(s) = \frac{Y_R}{Y_R + Y_L + Y_C} = \frac{1/R}{(1/R) + (1/sL) + sC} \tag{17.39}$$

$$= \frac{s(1/CR)}{s^2 + s(1/CR) + (1/LC)}$$

17.5.6 Realization of the Notch Functions

To obtain a pair of transmission zeros on the $j\omega$ axis, we use a parallel resonance circuit in the series arm, as shown in Fig. 17.18(e). Observe that this circuit is obtained by disconnecting both nodes x and y from ground and connecting them together to V_i. The impedance of the LC circuit becomes infinite at $\omega = \omega_0 = 1/\sqrt{LC}$, thus causing zero transmission at this frequency. The shunt impedance is resistive and thus does not introduce transmission zeros. It follows

that the circuit in Fig. 17.18(e) will realize the notch transfer function

$$T(s) = a_2 \frac{s^2 + \omega_0^2}{s^2 + s(\omega_0/Q) + \omega_0^2} \tag{17.40}$$

The value of the high-frequency gain a_2 can be found from the circuit to be unity.

To obtain a notch-filter realization in which the notch frequency ω_n is arbitrarily placed relative to ω_0, we adopt a variation on the scheme above. We still use a parallel LC circuit in the series branch, as shown in Fig. 17.18(f), where L_1 and C_1 are selected so that

$$L_1 C_1 = 1/\omega_n^2 \tag{17.41}$$

Thus the L_1C_1 tank circuit will introduce a pair of transmission zeros at $\pm j\omega_n$, provided the L_2C_2 tank is not resonant at ω_n. Apart from this restriction, the values of L_2 and C_2 must be selected to ensure that the natural modes have not been altered; thus,

$$C_1 + C_2 = C \tag{17.42}$$

$$L_1 \| L_2 = L \tag{17.43}$$

In other words, when V_i is replaced by a short circuit, the circuit should reduce to the original LCR resonator. Another way of thinking about the circuit of Fig. 17.18(f) is that it is obtained from the original LCR resonator by lifting part of L and part of C off ground and connecting them to V_i.

It should be noted that in the circuit of Fig. 17.18(f), L_2 does *not* introduce a zero at $s = 0$ because at $s = 0$, the L_1C_1 circuit also has a zero. In fact, at $s = 0$ the circuit reduces to an inductive voltage divider with the dc transmission being $L_2/(L_1 + L_2)$. Similar comments can be made about C_2 and the fact that it does *not* introduce a zero at $s = \infty$ and that the transmission at $s = \infty$ is $C_1/(C_1 + C_2)$.

The LPN and HPN filter realizations are special cases of the general notch circuit of Fig. 17.18(f). Specifically, for the LPN,

$$\omega_n > \omega_0$$

and thus

$$L_1 C_1 < (L_1 \| L_2)(C_1 + C_2)$$

This condition can be satisfied with L_2 eliminated (i.e., $L_2 = \infty$ and $L_1 = L$), resulting in the LPN circuit in Fig. 17.18(g). The transfer function can be written by inspection as

$$T(s) \equiv \frac{V_o}{V_i} = a_2 \frac{s^2 + \omega_n^2}{s^2 + s(\omega_0/Q) + \omega_0^2} \tag{17.44}$$

where $\omega_n^2 = 1/LC_1$, $\omega_0^2 = 1/L(C_1 + C_2)$, $\omega_0/Q = 1/CR$, and a_2 is the high-frequency gain. From the circuit we see that as $s \to \infty$, the circuit reduces to that in Fig. 17.18(h), for which

$$\frac{V_o}{V_i} = \frac{C_1}{C_1 + C_2}$$

Thus,

$$a_2 = \frac{C_1}{C_1 + C_2} \tag{17.45}$$

To obtain an HPN realization we start with the circuit of Fig. 17.18(f) and use the fact that $\omega_n < \omega_0$ to obtain

$$L_1 C_1 > (L_1 \| L_2)(C_1 + C_2)$$

which can be satisfied while selecting $C_2 = 0$ (i.e., $C_1 = C$). Thus we obtain the reduced circuit shown in Fig. 17.18(i). Observe that as $s \to \infty$, V_o approaches V_i and thus the high-frequency gain is unity. Thus, the transfer function can be expressed as

$$T(s) \equiv \frac{V_o}{V_i} = \frac{s^2 + (1/L_1 C)}{s^2 + s(1/CR) + [1/(L_1 \| L_2)C]} \tag{17.46}$$

17.5.7 Realization of the All-Pass Function

The all-pass transfer function

$$T(s) = \frac{s^2 - s(\omega_0/Q) + \omega_0^2}{s^2 + s(\omega_0/Q) + \omega_0^2} \tag{17.47}$$

can be written as

$$T(s) = 1 - \frac{s2(\omega_0/Q)}{s^2 + s(\omega_0/Q) + \omega_0^2} \tag{17.48}$$

The second term on the right-hand side is a bandpass function with a center-frequency gain of 2. We already have a bandpass circuit (Fig. 17.18d), but with a center-frequency gain of unity. We shall therefore attempt an all-pass realization with a flat gain of 0.5, that is,

$$T(s) = 0.5 - \frac{s(\omega_0/Q)}{s^2 + s(\omega_0/Q) + \omega_0^2}$$

This function can be realized using a voltage divider with a transmission ratio of 0.5 together with the bandpass circuit of Fig. 17.18(d). To effect the subtraction, the output of the all-pass circuit is taken between the output terminal of the voltage divider and that of the bandpass filter, as shown in Fig. 17.19. Unfortunately this circuit has the disadvantage of lacking a common ground terminal between the input and the output. An op amp–RC realization of the all-pass function will be presented in the next section.

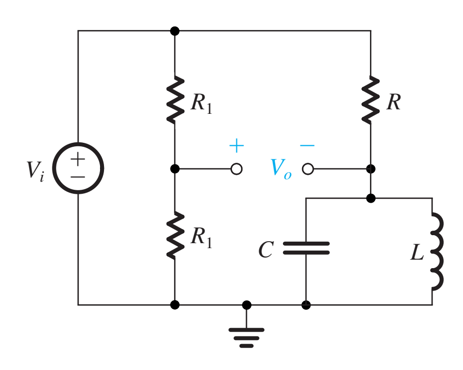


Figure 17.19 Realization of the second-order all-pass transfer function using a voltage divider and an LCR resonator.

EXERCISES

17.17 Use the circuit of Fig. 17.18(b) to realize a second-order low-pass function of the maximally flat type with a 3-dB frequency of 100 kHz.
Ans. Selecting $R = 1\ \text{k}\Omega$, we obtain $C = 1125$ pF and $L = 2.25$ mH.

17.18 Use the circuit of Fig. 17.18(e) to design a notch filter to eliminate a bothersome power-supply hum at a 60-Hz frequency. The filter is to have a 3-dB bandwidth of 10 Hz (i.e., the attenuation is greater than 3 dB over a 10-Hz band around the 60-Hz center frequency; see Exercise 17.15 and Fig. 17.16d). Use $R = 10\ \text{k}\Omega$.
Ans. $C = 1.6\ \mu\text{F}$ and $L = 4.42$ H (Note the large inductor required. This is why passive filters are not practical in low-frequency applications.)

17.6 Second-Order Active Filters Based on Inductor Replacement

In this section, we study a family of op amp–RC circuits that realize the various second-order filter functions. The circuits are based on an op amp–RC resonator obtained by replacing the inductor L in the LCR resonator with an op amp–RC circuit that has an inductive input impedance.

17.6.1 The Antoniou Inductance-Simulation Circuit

Over the years, many op amp–RC circuits have been proposed for simulating the operation of an inductor. Of these, one circuit invented by A. Antoniou[6] (see Antoniou, 1969) has proved to be the "best." By "best" we mean that the operation of the circuit is very tolerant of the nonideal properties of the op amps, in particular their finite gain and bandwidth. Figure 17.20(a) shows the Antoniou inductance-simulation circuit. If the circuit is fed at its input (node 1) with a voltage source V_1 and the input current is denoted I_1, then for ideal op amps the input impedance can be shown to be

$$Z_{in} \equiv V_1/I_1 = sC_4R_1R_3R_5/R_2 \tag{17.49}$$

which is that of an inductance L given by



$$L = C_4R_1R_3R_5/R_2 \tag{17.50}$$

Figure 17.20(b) shows the analysis of the circuit assuming that the op amps are ideal and thus that a virtual short circuit appears between the two input terminals of each op amp, and assuming also that the input currents of the op amps are zero. The analysis begins at node 1, which is assumed to be fed by a voltage source V_1, and proceeds step by step, with the order of the steps indicated by the circled numbers. The result of the analysis is the expression shown for the input current I_1 from which Z_{in} is found.

[6] Andreas Antoniou is a Canadian academic, currently (2014) a professor emeritus of electrical and computer engineering at the University of Victoria, Victoria, British Columbia.

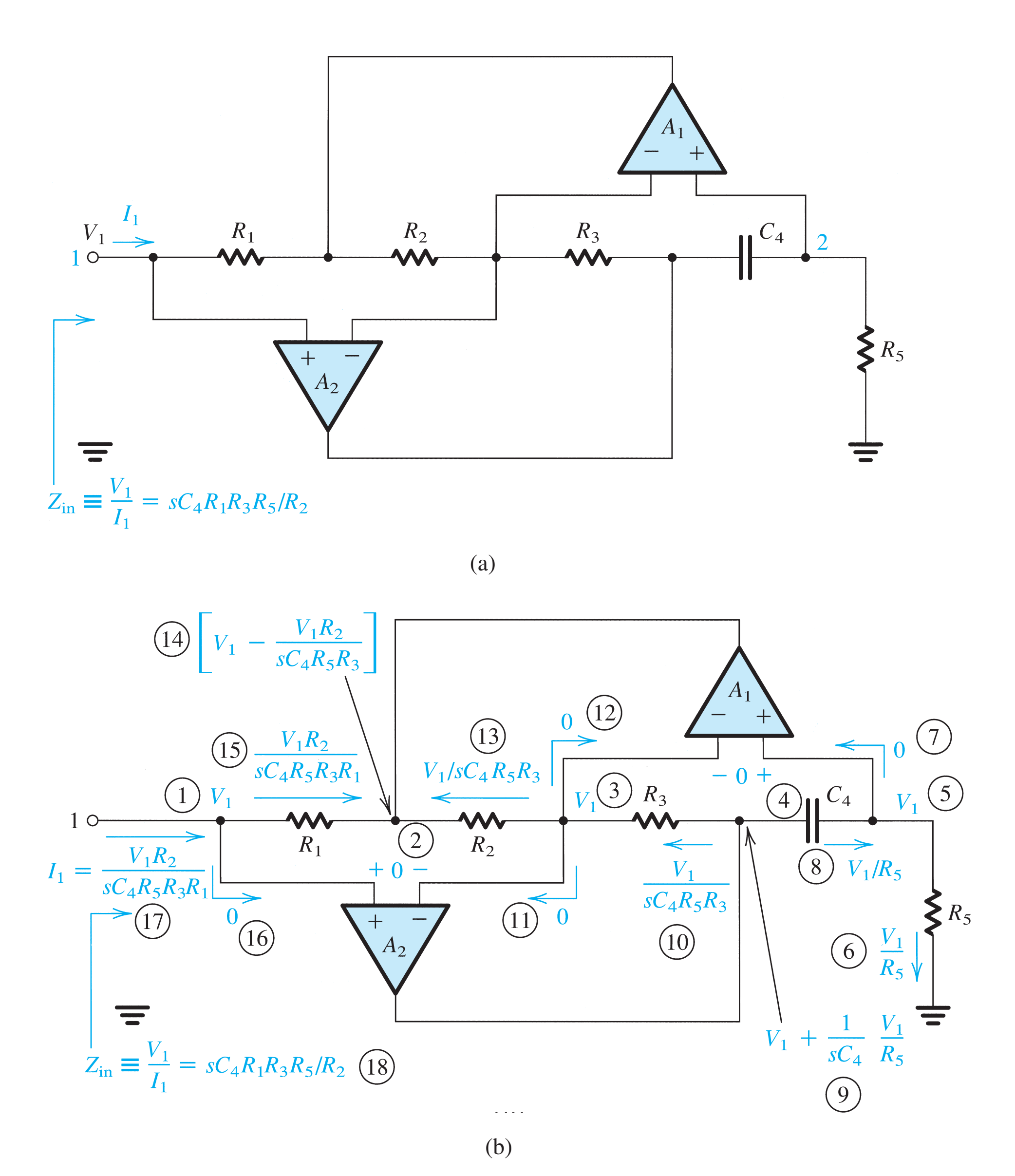
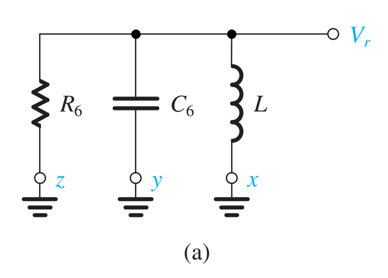


Figure 17.20 **(a)** The Antoniou inductance-simulation circuit. **(b)** Analysis of the circuit assuming ideal op amps. The order of the analysis steps is indicated by the circled numbers.

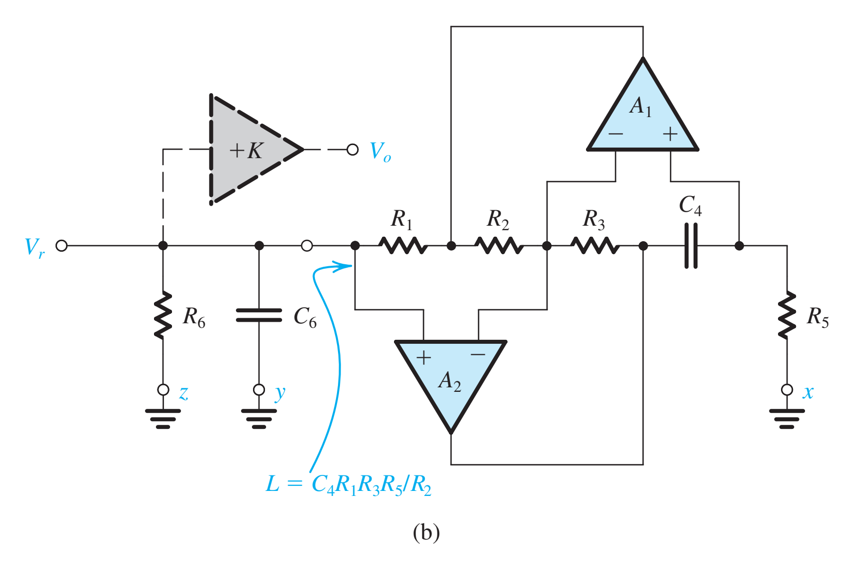
The design of this circuit is usually based on selecting $R_1 = R_2 = R_3 = R_5 = R$ and $C_4 = C$, which leads to $L = CR^2$. Convenient values are then selected for C and R to yield the desired inductance value L. More details on this circuit and the effect of the nonidealities of the op amps on its performance can be found in Sedra and Brackett (1978).

17.6.2 The Op Amp–RC Resonator

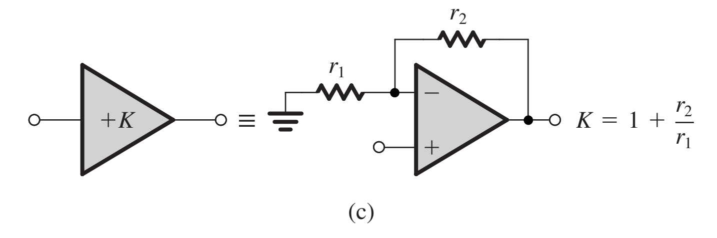
Figure 17.21(a) shows the LCR resonator we studied in detail in Section 17.5. Replacing the inductor L with a simulated inductance realized by the Antoniou circuit of Fig. 17.20(a) results in the op amp–RC resonator of Fig. 17.21(b). (Ignore for the moment the additional amplifier



(a)



(b)



(c)

Figure 17.21 **(a)** An LCR resonator. **(b)** An op amp–RC resonator obtained by replacing the inductor L in the LCR resonator of **(a)** with a simulated inductance realized by the Antoniou circuit of Fig. 17.20(a). **(c)** Implementation of the buffer amplifier K.

drawn with broken lines.) The circuit of Fig. 17.21(b) is a second-order resonator having a pole frequency

$$\omega_0 = 1/\sqrt{LC_6} = 1/\sqrt{C_4C_6R_1R_3R_5/R_2} \tag{17.51}$$

where we have used the expression for L given in Eq. (17.50). The pole Q factor can be obtained using the expression in Eq. (17.35) with $C = C_6$ and $R = R_6$; thus, $Q = \omega_0\, C_6\, R_6$.

Replacing ω_0 by the expression in Eq. (17.51) gives

$$Q = \omega_0 C_6 R_6 = R_6 \sqrt{\frac{C_6}{C_4} \frac{R_2}{R_1 R_3 R_5}} \tag{17.52}$$

Usually one selects $C_4 = C_6 = C$ and $R_1 = R_2 = R_3 = R_5 = R$, which results in

$$\omega_0 = 1/CR \tag{17.53}$$
$$Q = R_6/R \tag{17.54}$$

Thus, if we select a practically convenient value for C, we can use Eq. (17.53) to determine the value of R to realize a given ω_0, and then use Eq. (17.54) to determine the value of R_6 to realize a given Q.

17.6.3 Realization of the Various Filter Types

The op amp–RC resonator of Fig. 17.21(b) can be used to generate circuit realizations for the various second-order filter functions by following the approach described in detail in Section 17.5 in connection with the LCR resonator. Thus to obtain a bandpass function, we disconnect node z from ground and connect it to the signal source V_i. A high-pass function is obtained by injecting V_i to node y. To realize a low-pass function using the LCR resonator, the inductor terminal x is disconnected from ground and connected to V_i. The corresponding node in the active resonator is the node at which R_5 is connected to ground,[7] labeled as node x in Fig. 17.21(b). A regular notch function ($\omega_n = \omega_0$) is obtained by feeding V_i to nodes x and y. In all cases the output can be taken as the voltage across the resonance circuit, V_r. However, this is not a convenient node to use as the filter output terminal because connecting a load there would change the filter characteristics. The problem can be solved easily by utilizing a buffer amplifier. This is the amplifier of gain K, drawn with broken lines in Fig. 17.21(b). Figure 17.21(c) shows how this amplifier can be simply implemented using an op amp connected in the noninverting configuration. Note that not only does the amplifier K buffer the output of the filter, but it also allows the designer to set the filter gain to any desired value by appropriately selecting the value of K.

Figure 17.22 shows the various second-order filter circuits obtained from the resonator of Fig. 17.21(b). The transfer functions and design equations for these circuits are given in Table 17.1. Note that the transfer functions can be written by analogy to those of the LCR resonator. We have already commented on the LP, HP, BP, and regular-notch circuits given in Fig. 17.22(a) to (d). The LPN and HPN circuits in Fig. 17.22(e) and (f) are obtained by direct analogy to their LCR counterparts in Fig. 17.18(g) and (i), respectively. The all-pass circuit in Fig. 17.22(g), however, deserves some explanation.

17.6.4 The All-Pass Circuit

From Eq. (17.48) we see that an all-pass function with a flat gain of unity can be written as

$$\text{AP} = 1 - (\text{BP with a center-frequency gain of 2}) \tag{17.55}$$

Two circuits whose transfer functions are related in this fashion (i.e., the transfer function of one is equal to unity minus the transfer function of the other) are said to be

[7]This point might not be obvious! The reader, however, can show by direct analysis that when V_i is fed to this node, the function V_r/V_i is indeed low pass.

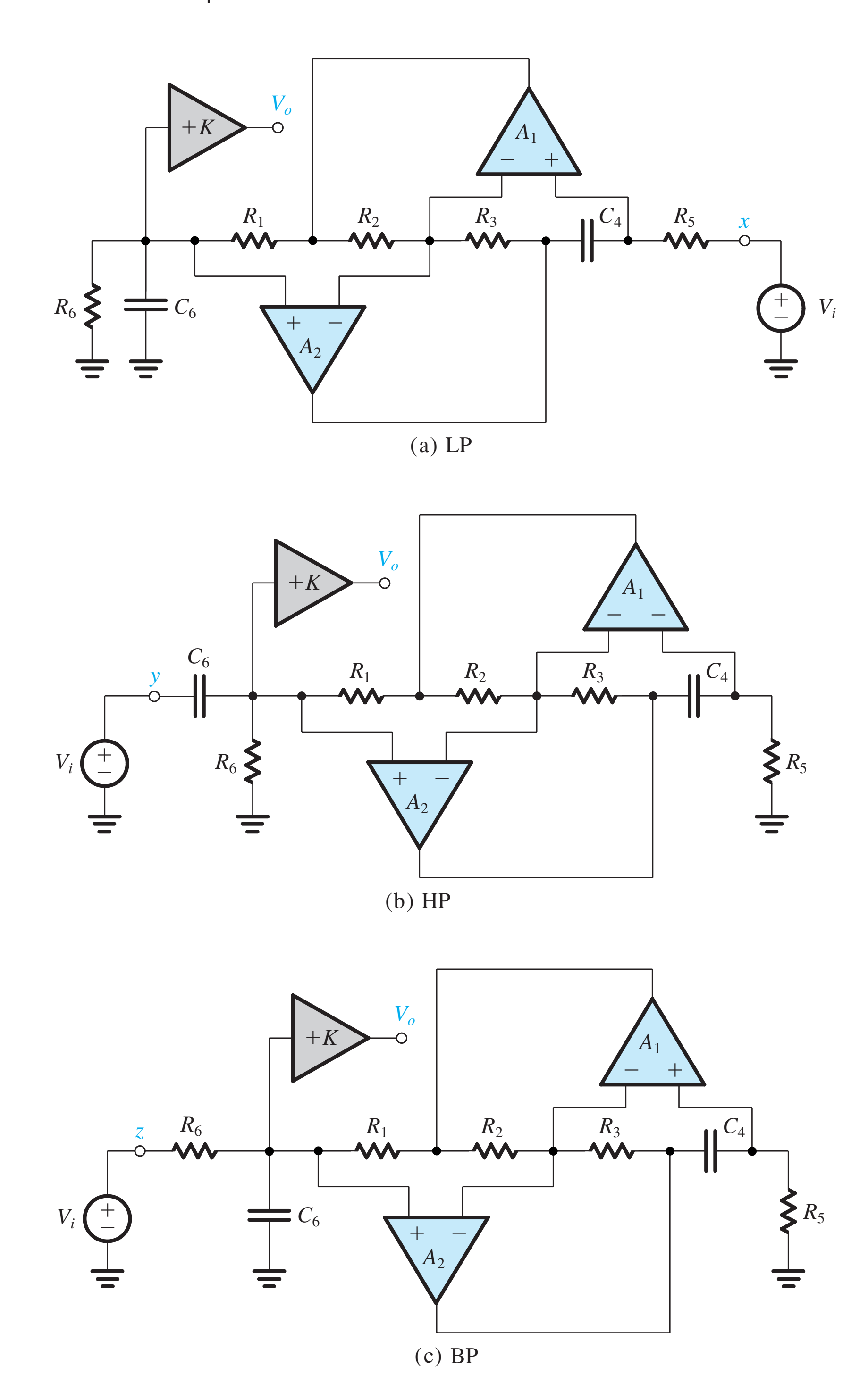


Figure 17.22 Realizations for the various second-order filter functions using the op amp–RC resonator of Fig. 17.21(b): **(a)** LP, **(b)** HP, **(c)** BP. The circuits are based on the LCR circuit in Fig. 17.18. Design considerations are given in Table 17.1.

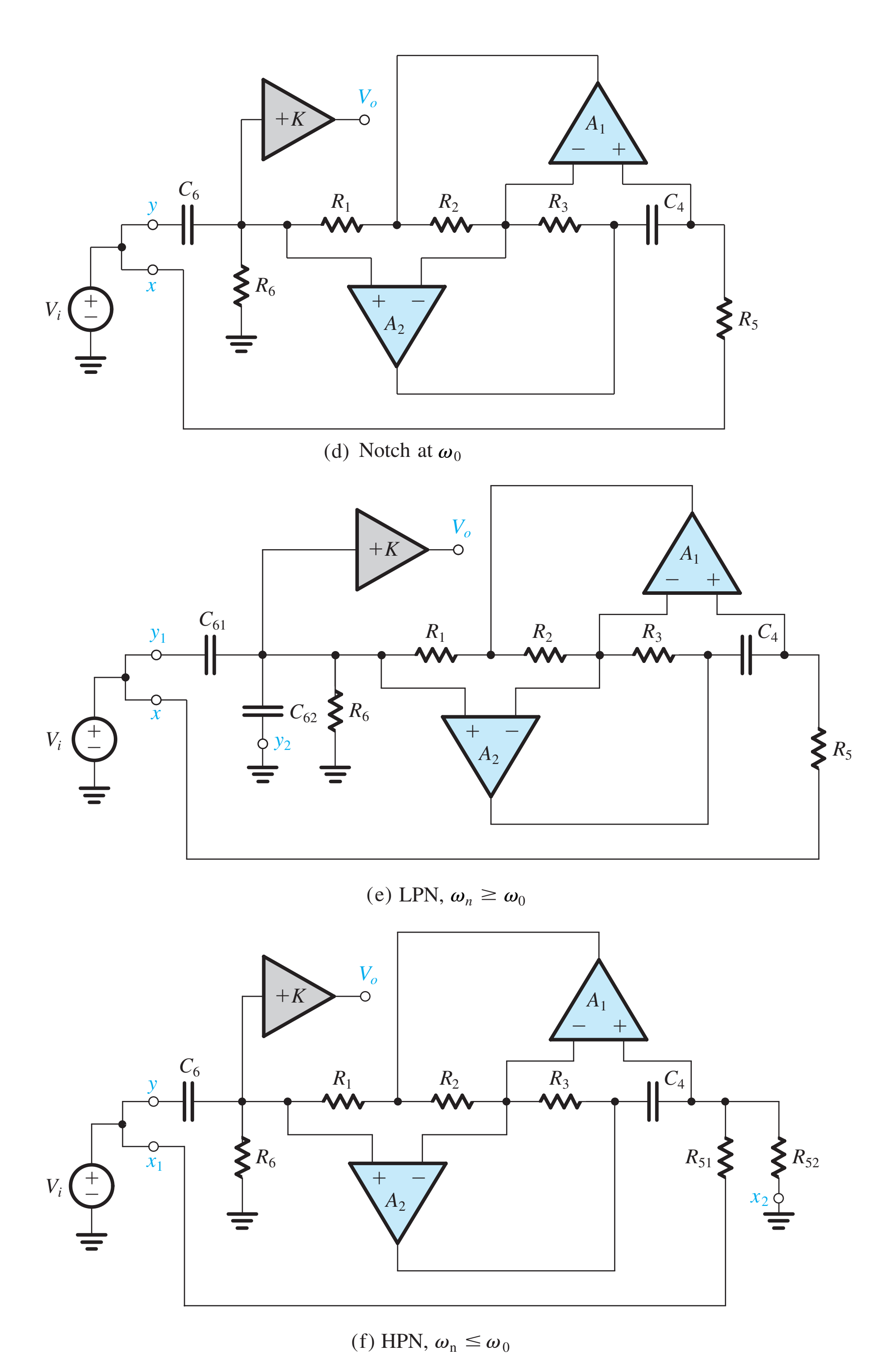
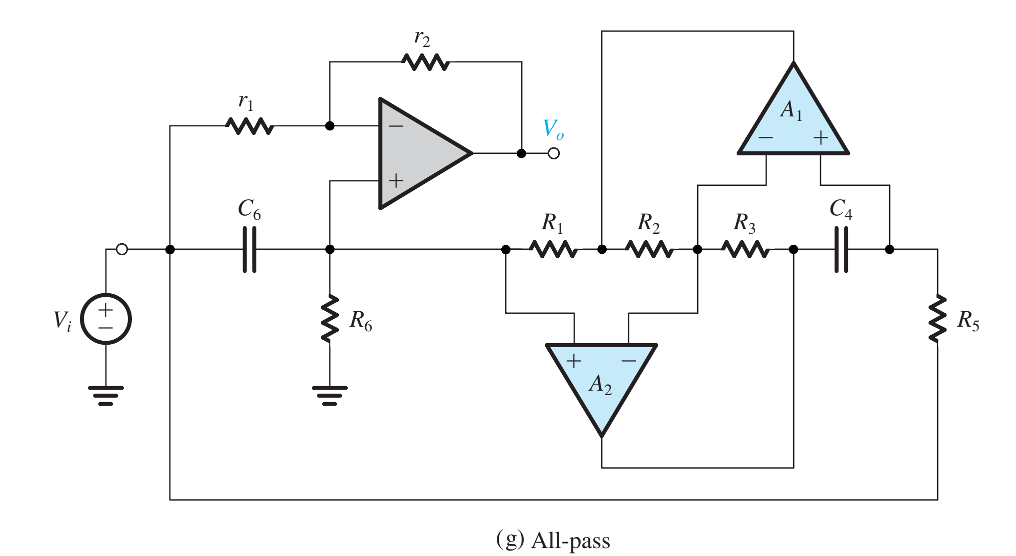


Figure 17.22 *continued* **(d)** Notch at ω_0; **(e)** LPN, $\omega_n \geq \omega_0$; **(f)** HPN, $\omega_n \leq \omega_0$.



(g) All-pass

Figure 17.22 *continued* **(g)** All pass.

Table 17.1 Design Data for the Circuits of Fig. 17.22

Circuit	Transfer Function and Other Parameters	Design Equations
Resonator Fig. 17.21(b)	$\omega_0 = 1/\sqrt{C_4 C_6 R_1 R_3 R_5 / R_2}$ $Q = R_6 \sqrt{\frac{C_6}{C_4} \frac{R_2}{R_1 R_3 R_5}}$	$C_4 = C_6 = C$ (practical value) $R_1 = R_2 = R_3 = R_5 = 1/\omega_0 C$ $R_6 = Q/\omega_0 C$
Low-pass (LP) Fig. 17.22(a)	$T(s) = \frac{KR_2/C_4 C_6 R_1 R_3 R_5}{s^2 + s\frac{1}{C_6 R_6} + \frac{R_2}{C_4 C_6 R_1 R_3 R_5}}$	$K =$ DC gain
High-pass (HP) Fig. 17.22(b)	$T(s) = \frac{Ks^2}{s^2 + s\frac{1}{C_6 R_6} + \frac{R_2}{C_4 C_6 R_1 R_3 R_5}}$	$K =$ High-frequency gain
Bandpass (BP) Fig. 17.22(c)	$T(s) = \frac{Ks/C_6 R_6}{s^2 + s\frac{1}{C_6 R_6} + \frac{R_2}{C_4 C_6 R_1 R_3 R_5}}$	$K =$ Center-frequency gain
Regular notch (N) Fig. 17.22(d)	$T(s) = \frac{K\left[s^2 + (R_2/C_4 C_6 R_1 R_3 R_5)\right]}{s^2 + s\frac{1}{C_6 R_6} + \frac{R_2}{C_4 C_6 R_1 R_3 R_5}}$	$K =$ Low- and high-frequency gain

Low-pass notch (LPN) Fig. 17.22(e)	$T(s) = K\frac{C_{61}}{C_{61}+C_{62}}$ $\times \frac{s^2 + (R_2/C_4C_{61}R_1R_3R_5)}{s^2 + s\frac{1}{(C_{61}+C_{62})R_6} + \frac{R_2}{C_4(C_{61}+C_{62})R_1R_3R_5}}$ $\omega_n = 1/\sqrt{C_4C_{61}R_1R_3R_5/R_2}$ $\omega_0 = 1/\sqrt{C_4(C_{61}+C_{62})R_1R_3R_5/R_2}$ $Q = R_6\sqrt{\frac{C_{61}+C_{62}}{C_4}\frac{R_2}{R_1R_3R_5}}$	$K =$ DC gain $C_{61}+C_{62} = C_6 = C$ $C_{61} = C(\omega_0/\omega_n)^2$ $C_{62} = C - C_{61}$
High-pass notch (HPN) Fig. 17.22(f)	$T(s) = K\frac{s^2 + (R_2/C_4C_6R_1R_3R_{51})}{s^2 + s\frac{1}{C_6R_6} + \frac{R_2}{C_4C_6R_1R_3}\left(\frac{1}{R_{51}}+\frac{1}{R_{52}}\right)}$ $\omega_n = 1/\sqrt{C_4C_6R_1R_3R_{51}/R_2}$ $\omega_0 = \sqrt{\frac{R_2}{C_4C_6R_1R_3}\left(\frac{1}{R_{51}}+\frac{1}{R_{52}}\right)}$ $Q = R_6\sqrt{\frac{C_6}{C_4}\frac{R_2}{R_1R_3}\left(\frac{1}{R_{51}}+\frac{1}{R_{52}}\right)}$	$K =$ High-frequency gain $\frac{1}{R_{51}}+\frac{1}{R_{52}} = \frac{1}{R_5} = \omega_0 C$ $R_{51} = R_5(\omega_0/\omega_n)^2$ $R_{52} = R_5\Big/\left[1-(\omega_n/\omega_0)^2\right]$
All-pass (AP) Fig. 17.22(g)	$T(s) = \frac{s^2 - s\frac{1}{C_6R_6}\frac{r_2}{r_1} + \frac{R_2}{C_4C_6R_1R_3R_5}}{s^2 + s\frac{1}{C_6R_6} + \frac{R_2}{C_4C_6R_1R_3R_5}}$ $\omega_z = \omega_0$ $Q_z = Q(r_1/r_2)$ Flat gain $= 1$	$r_1 = r_2 = r$ (arbitrary) Adjust r_2 to make $Q_z = Q$

complementary.[8] Thus the all-pass circuit with unity flat gain is the complement of the bandpass circuit with a center-frequency gain of 2. A simple procedure exists for obtaining the complement of a given linear circuit: Disconnect all the circuit nodes that are connected to ground and connect them to V_i, and disconnect all the nodes that are connected to V_i and connect them to ground. That is, *interchanging input and ground in a linear circuit generates a circuit whose transfer function is the complement of that of the original circuit*.

Returning to the problem at hand, we first use the circuit of Fig. 17.22(c) to realize a BP with a gain of 2 by simply selecting $K = 2$ and implementing the buffer amplifier with the circuit of Fig. 17.21(c) with $r_1 = r_2$. We then interchange input and ground and thus obtain the all-pass circuit of Fig. 17.22(g).

Finally, in addition to being simple to design, the circuits in Fig. 17.22 exhibit excellent performance. They can be used on their own to realize second-order filter functions, or they can be cascaded to implement high-order filters.

[8]More about complementary circuits will be presented later in conjunction with Fig. 17.31.

EXERCISES

D17.19 Use the circuit of Fig. 17.22(c) to design a second-order bandpass filter with a center frequency of 10 kHz, a 3-dB bandwidth of 500 Hz, and a center-frequency gain of 10. Use $C = 1.2$ nF.
Ans. $R_1 = R_2 = R_3 = R_5 = 13.26$ kΩ; $R_6 = 265$ kΩ; $C_4 = C_6 = 1.2$ nF; $K = 10$, $r_1 = 10$ kΩ, $r_2 = 90$ kΩ

D17.20 Realize the Chebyshev filter of Example 17.2, whose transfer function is given in Eq. (17.25), as the cascade connection of three circuits: two of the type shown in Fig. 17.22(a) and one first-order op amp–RC circuit of the type shown in Fig. 17.13(a). Note that you can make the dc gain of all sections equal to unity. Do so. Use as many 10-kΩ resistors as possible.
Ans. First-order section: $R_1 = R_2 = 10$ kΩ, $C = 5.5$ nF; second-order section with $\omega_0 = 4.117 \times 10^4$ rad/s and $Q = 1.4$: $R_1 = R_2 = R_3 = R_5 = 10$ kΩ, $R_6 = 14$ kΩ, $C_4 = C_6 = 2.43$ nF, $r_1 = \infty$, $r_2 = 0$; second-order section with $\omega_0 = 6.246 \times 10^4$ rad/s and $Q = 5.56$: $R_1 = R_2 = R_3 = R_5 = 10$ kΩ, $R_6 = 55.6$ kΩ, $C_4 = C_6 = 1.6$ nF, $r_1 = \infty$, $r_2 = 0$

17.7 Second-Order Active Filters Based on the Two-Integrator-Loop Topology

In this section, we study another family of op amp–RC circuits that realize second-order filter functions. The circuits are based on the use of two integrators connected in cascade in an overall feedback loop and are thus known as two-integrator-loop circuits.

17.7.1 Derivation of the Two-Integrator-Loop Biquad

To derive the two-integrator-loop biquadratic circuit, or **biquad** as it is commonly known,[9] consider the second-order high-pass transfer function

$$\frac{V_{hp}}{V_i} = \frac{Ks^2}{s^2 + s(\omega_0/Q) + \omega_0^2} \tag{17.56}$$

where K is the high-frequency gain. Cross-multiplying Eq. (17.56) and dividing both sides of the resulting equation by s^2 (to get all the terms involving s in the form $1/s$, which is the transfer function of an integrator) gives

$$V_{hp} + \frac{1}{Q}\left(\frac{\omega_0}{s}V_{hp}\right) + \left(\frac{\omega_0^2}{s^2}V_{hp}\right) = KV_i \tag{17.57}$$

In this equation we observe that the signal $(\omega_0/s)V_{hp}$ can be obtained by passing V_{hp} through an integrator with a time constant equal to $1/\omega_0$. Furthermore, passing the resulting signal through

[9]The name *biquad* stems from the fact that this circuit in its most general form is capable of realizing a biquadratic transfer function, that is, one that is the ratio of two quadratic polynomials.

another identical integrator results in the third signal involving V_{hp} in Eq. (17.57)—namely, $(\omega_0^2/s^2)V_{hp}$. Figure 17.23(a) shows a block diagram for such a two-integrator arrangement. Note that in anticipation of the use of the inverting op-amp Miller integrator circuit (Section 2.5.2) to implement each integrator, the integrator blocks in Fig. 17.23(a) have been assigned negative signs.

The problem still remains, however, of how to form V_{hp}, the input signal feeding the two cascaded integrators. Toward that end, we rearrange Eq. (17.57), expressing V_{hp} in terms of its single- and double-integrated versions and of V_i as

$$V_{hp} = KV_i - \frac{1}{Q}\frac{\omega_0}{s}V_{hp} - \frac{\omega_0^2}{s^2}V_{hp} \tag{17.58}$$

which suggests that V_{hp} can be obtained by using the weighted summer of Fig. 17.23(b). Now it should be easy to see that a complete block diagram realization can be obtained by combining the integrator blocks of Fig. 17.23(a) with the summer block of Fig. 17.23(b), as shown in Fig. 17.23(c).

In the realization of Fig. 17.23(c), V_{hp}, obtained at the output of the summer, realizes the high-pass transfer function $T_{hp} \equiv V_{hp}/V_i$ of Eq. (17.56). The signal at the output of the first integrator is $-(\omega_0/s)V_{hp}$, which is a bandpass function,

$$\frac{(-\omega_0/s)V_{hp}}{V_i} = -\frac{K\omega_0 s}{s^2 + s(\omega_0/Q) + \omega_0^2} = T_{bp}(s) \tag{17.59}$$

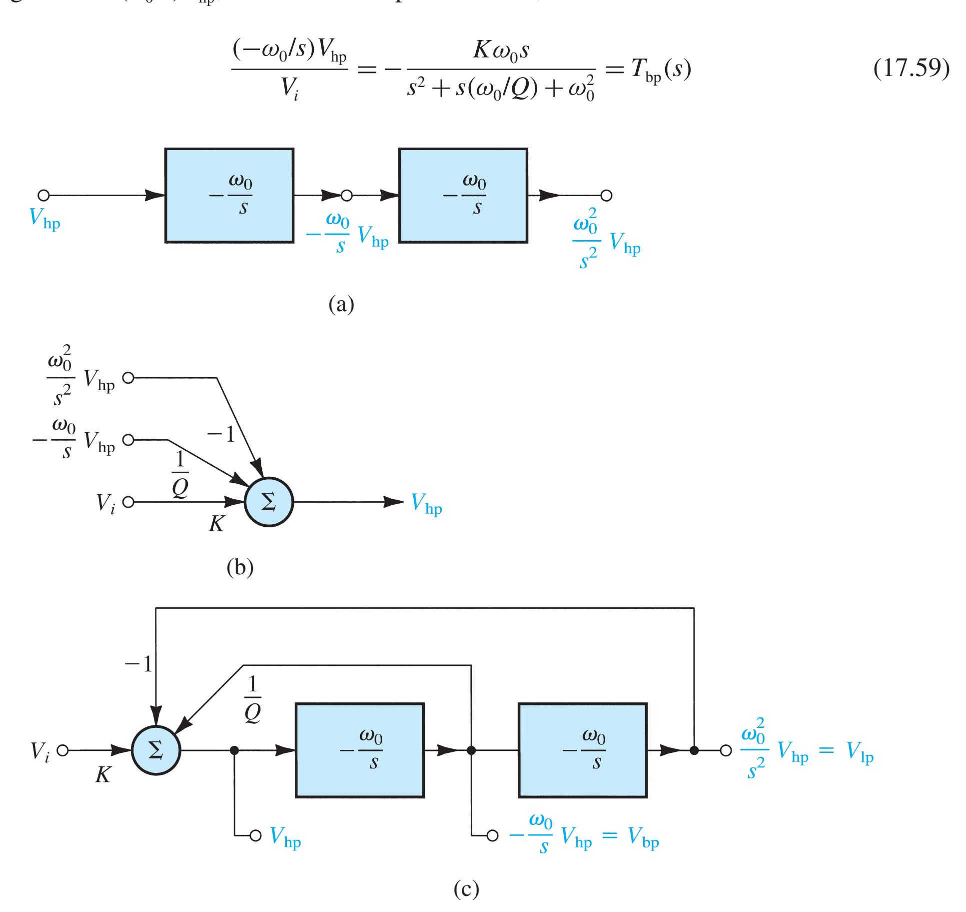


Figure 17.23 Derivation of a block diagram realization of the two-integrator-loop biquad.

Therefore the signal at the output of the first integrator is labeled V_{bp}. Note that the center-frequency gain of the bandpass filter realized is equal to $-KQ$.

In a similar fashion, we can show that the transfer function realized at the output of the second integrator is the low-pass function,

$$\frac{(\omega_0^2/s^2)V_{hp}}{V_i} = \frac{K\omega_0^2}{s^2 + s(\omega_0/Q) + \omega_0^2} = T_{lp}(s) \tag{17.60}$$

Thus the output of the second integrator is labeled V_{lp}. Note that the dc gain of the low-pass filter realized is equal to K.

We conclude that the two-integrator-loop biquad shown in block diagram form in Fig. 17.23(c) realizes the three basic second-order filtering functions, LP, BP, and HP, *simultaneously*. This versatility has made the circuit very popular and has given it the name *universal active filter*.

17.7.2 Circuit Implementation

To obtain an op-amp circuit implementation of the two-integrator-loop biquad of Fig. 17.23(c), we replace each integrator with a Miller integrator circuit having $CR = 1/\omega_0$, and we replace the summer block with an op-amp summing circuit that is capable of assigning both positive and negative weights to its inputs. The resulting circuit, known as the Kerwin–Huelsman–Newcomb or **KHN biquad**, after its inventors, is shown in Fig. 17.24(a). Given values for ω_0, Q, and K, the design of the circuit is straightforward: We select suitably practical values for the components C and R of the integrators so that $CR = 1/\omega_0$. To determine the values of the resistors associated with the summer, we first use *superposition* to express the output of the summer V_{hp} in terms of its inputs, V_i, V_{bp}, and V_{lp} as

$$V_{hp} = V_i\frac{R_3}{R_2+R_3}\left(1+\frac{R_f}{R_1}\right) + V_{bp}\frac{R_2}{R_2+R_3}\left(1+\frac{R_f}{R_1}\right) - V_{lp}\frac{R_f}{R_1}$$

Substituting $V_{bp} = -(\omega_0/s)V_{hp}$ and $V_{lp} = ({\omega_0}^2/s^2)V_{hp}$ gives

$$V_{hp} = \frac{R_3}{R_2+R_3}\left(1+\frac{R_f}{R_1}\right)V_i + \frac{R_2}{R_2+R_3}\left(1+\frac{R_f}{R_1}\right)\left(-\frac{\omega_0}{s}V_{hp}\right) - \frac{R_f}{R_1}\left(\frac{\omega_0^2}{s^2}V_{hp}\right) \tag{17.61}$$

Equating the last right-hand-side terms of Eqs. (17.61) and (17.58) gives

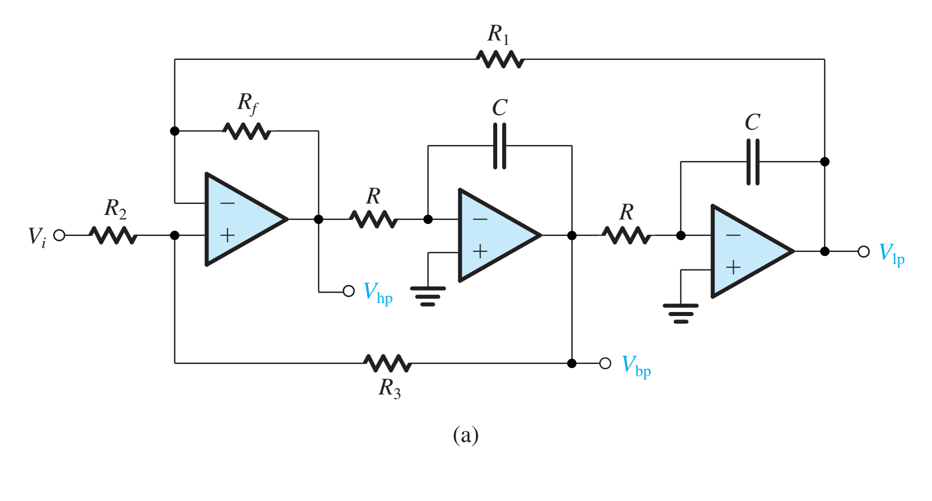
$$R_f/R_1 = 1 \tag{17.62}$$

which implies that we can select arbitrary but practically convenient equal values for R_1 and R_f. Then, equating the second-to-last terms on the right-hand side of Eqs. (17.61) and (17.58) and setting $R_1 = R_f$ yields the ratio R_3/R_2 required to realize a given Q as

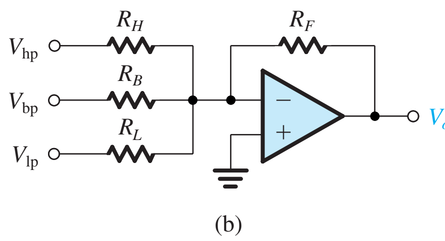
$$R_3/R_2 = 2Q - 1 \tag{17.63}$$

Thus an arbitrary but convenient value can be selected for either R_2 or R_3, and the value of the other resistance can be determined using Eq. (17.63). Finally, equating the coefficients of V_i in Eqs. (17.61) and (17.58) and substituting $R_f = R_1$ and for R_3/R_2 from Eq. (17.63) results in

$$K = 2 - (1/Q) \tag{17.64}$$



(a)



(b)

Figure 17.24 **(a)** The KHN biquad circuit, obtained as a direct implementation of the block diagram of Fig. 17.23(c). The three basic filtering functions, HP, BP, and LP, are simultaneously realized. **(b)** To obtain notch and all-pass functions, the three outputs are summed with appropriate weights using this op-amp summer.

Thus the gain parameter K is fixed to this value.

The KHN biquad can be used to realize notch and all-pass functions by summing weighted versions of the three outputs, LP, BP, and HP. Such an op-amp summer is shown in Fig. 17.24(b); for this summer we can write

$$V_o = -\left(\frac{R_F}{R_H}V_{\text{hp}} + \frac{R_F}{R_B}V_{\text{bp}} + \frac{R_F}{R_L}V_{\text{lp}}\right) \tag{17.65}$$

$$= -V_i\left(\frac{R_F}{R_H}T_{\text{hp}} + \frac{R_F}{R_B}T_{\text{bp}} + \frac{R_F}{R_L}T_{\text{lp}}\right)$$

Substituting for T_{hp}, T_{bp}, and T_{lp} from Eqs. (17.56), (17.59), and (17.60), respectively, gives the overall transfer function

$$\frac{V_o}{V_i} = -K\frac{(R_F/R_H)s^2 - s(R_F/R_B)\omega_0 + (R_F/R_L)\omega_0^2}{s^2 + s(\omega_0/Q) + \omega_0^2} \tag{17.66}$$

from which we can see that different transmission zeros can be obtained by the appropriate selection of the values of the summing resistors. For instance, a notch is obtained by selecting $R_B = \infty$ and

$$\frac{R_H}{R_L} = \left(\frac{\omega_n}{\omega_0}\right)^2 \tag{17.67}$$

17.7.3 An Alternative Two-Integrator-Loop Biquad Circuit

An alternative two-integrator-loop biquad circuit in which all three op amps are used in a single-ended mode can be developed as follows: Rather than using the input summer to add signals with positive and negative coefficients, we can introduce an additional inverter, as shown in Fig. 17.25(a). Now all the coefficients of the summer have the same sign, and we can dispense with the summing amplifier altogether and perform the summation at the virtual-ground input of the first integrator. Observe that the summing weights of 1, $1/Q$, and K are realized by using resistances of R, QR, and R/K, respectively. The resulting circuit is shown in Fig. 17.25(b), from which we observe that the high-pass function is no longer available! This is the price paid for obtaining a circuit that utilizes all op amps in a single-ended mode. The circuit of Fig. 17.25(b) is known as the **Tow–Thomas biquad**, after its originators.

Rather than using a fourth op amp to realize the finite transmission zeros required for the notch and all-pass functions, as was done with the KHN biquad, an economical *feedforward* scheme can be employed with the Tow–Thomas circuit. Specifically, the virtual ground available at the input of each of the three op amps in the Tow–Thomas circuit permits the input signal to be fed to all three op amps, as shown in Fig. 17.26. If V_o is taken at the output of the damped integrator, straightforward analysis yields the filter transfer function

$$\frac{V_o}{V_i} = -\frac{s^2\left(\frac{C_1}{C}\right) + s\frac{1}{C}\left(\frac{1}{R_1} - \frac{r}{RR_3}\right) + \frac{1}{C^2RR_2}}{s^2 + s\frac{1}{QCR} + \frac{1}{C^2R^2}} \tag{17.68}$$

which can be used to obtain the design data given in Table 17.2.

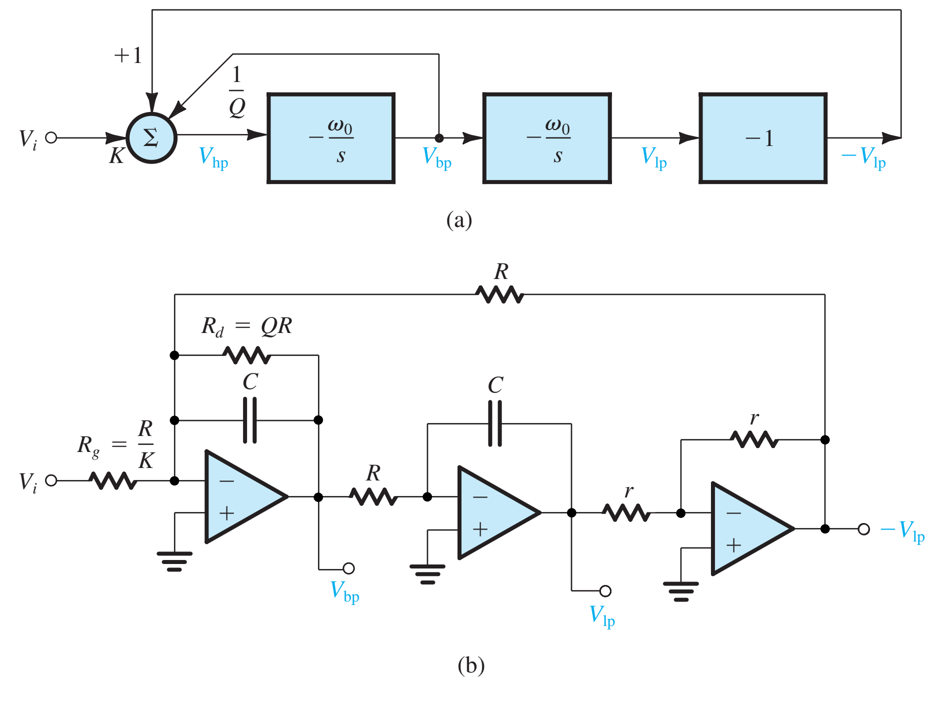


Figure 17.25 **(a)** Derivation of an alternative two-integrator-loop biquad in which all op amps are used in a single-ended fashion. **(b)** The resulting circuit, known as the Tow–Thomas biquad.

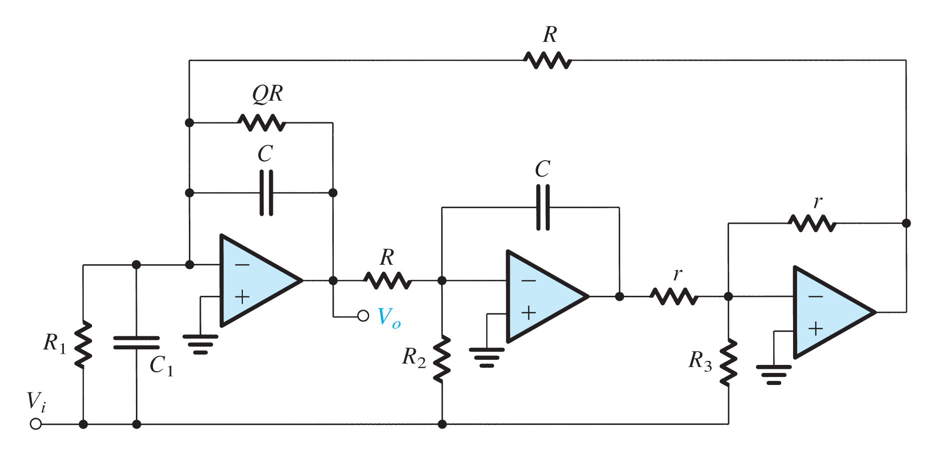


Figure 17.26 The Tow–Thomas biquad with feedforward. The transfer function of Eq. (17.68) is realized by feeding the input signal through appropriate components to the inputs of the three op amps. This circuit can realize all special second-order functions. The design equations are given in Table 17.2.

Table 17.2 Design Data for the Circuit in Fig. 17.26

All cases	$C =$ arbitrary, $R = 1/\omega_0 C$, $r =$ arbitrary
LP	$C_1 = 0$, $R_1 = \infty$, $R_2 = R/$dc gain, $R_3 = \infty$
Positive BP	$C_1 = 0$, $R_1 = \infty$, $R_2 = \infty$, $R_3 = Qr/$center-frequency gain
Negative BP	$C_1 = 0$, $R_1 = QR/$center-frequency gain, $R_2 = \infty$, $R_3 = \infty$
HP	$C_1 = C \times$ high-frequency gain, $R_1 = \infty$, $R_2 = \infty$, $R_3 = \infty$
Notch (all types)	$C_1 = C \times$ high-frequency gain, $R_1 = \infty$, $R_2 = R(\omega_0/\omega_n)^2/$high-frequency gain, $R_3 = \infty$
AP	$C_1 = C \times$ flat gain, $R_1 = \infty$, $R_2 = R/$gain, $R_3 = Qr/$gain

17.7.4 Final Remarks

Two-integrator-loop biquads are extremely versatile and easy to design. However, their performance is adversely affected by the finite bandwidth of the op amps. Special techniques exist for compensating the circuit for such effects (see the SPICE simulation example on the website, and Sedra and Brackett, 1978).

EXERCISES

D17.21 Design the KHN circuit to realize a high-pass function with $f_0 = 10$ kHz and $Q = 2$. Choose $C = 1$ nF. What is the value of high-frequency gain obtained? What is the center-frequency gain of the bandpass function that is simultaneously available at the output of the first integrator?
Ans. $R = 15.9$ kΩ; $R_1 = R_f = R_2 = 10$ kΩ (arbitrary); $R_3 = 30$ kΩ; 1.5; 3

D17.22 Use the KHN circuit together with an output summing amplifier to design a low-pass notch filter with $f_0 = 5$ kHz, $f_n = 8$ kHz, $Q = 5$, and a dc gain of 3. Select $C = 1$ nF and $R_L = 10$ kΩ.
Ans. $R = 31.83$ kΩ; $R_1 = R_f = R_2 = 10$ kΩ (arbitrary); $R_3 = 90$ kΩ; $R_H = 25.6$ kΩ; $R_F = 16.7$ kΩ; $R_B = \infty$

D17.23 Use the Tow–Thomas biquad [Fig. 17.25(b)] to design a second-order bandpass filter with $f_0 =$ 10 kHz, $Q = 20$, and unity center-frequency gain. If $R = 10$ kΩ, give the values of C, R_d, and R_g.
Ans. 1.59 nF; 200 kΩ; 200 kΩ

D17.24 Use the data of Table 17.2 to design the biquad circuit of Fig. 17.26 to realize an all-pass filter with $\omega_0 = 10^4$ rad/s, $Q = 5$, and flat gain = 1. Use $C = 10$ nF and $r = 10$ kΩ.
Ans. $R = 10$ kΩ; Q-determining resistor = 50 kΩ; $C_1 = 10$ nF; $R_1 = \infty$; $R_2 = 10$ kΩ; $R_3 = 50$ kΩ

17.8 Single-Amplifier Biquadratic Active Filters

The op amp–RC biquadratic circuits studied in the two preceding sections provide good performance, are versatile, and are easy to design and to adjust (tune) after final assembly. Unfortunately, however, they are not economic in their use of op amps, requiring three or four amplifiers per second-order section. This can be a problem, especially in applications that call for conservation of power-supply current: for instance, in a battery-operated instrument. In this section we shall study a class of second-order filter circuits that requires only one op amp per biquad. These minimal realizations, however, suffer a greater dependence on the limited gain and bandwidth of the op amp and can also be more sensitive to the unavoidable tolerances in the values of resistors and capacitors than the multiple-op-amp biquads of the preceding sections. The **single-amplifier biquads** (SABs) are therefore limited to the less stringent filter specifications—for example, pole Q factors less than about 10.

The synthesis of SAB circuits is based on the use of feedback to move the poles of an RC circuit from the negative real axis, where they naturally lie, to the complex-conjugate locations required to provide selective filter response. The synthesis of SABs follows a two-step process:

1. Synthesis of a feedback loop that realizes a pair of complex-conjugate poles characterized by a frequency ω_0 and a Q factor Q.
2. Injecting the input signal in a way that realizes the desired transmission zeros.

17.8.1 Synthesis of the Feedback Loop

Consider the circuit shown in Fig. 17.27(a), which consists of a two-port RC network n placed in the negative-feedback path of an op amp. We shall assume that, except for having a finite gain A, the op amp is ideal. We shall denote by $t(s)$ the open-circuit voltage-transfer function of the RC network n, where the definition of $t(s)$ is illustrated in Fig. 17.27(b). The transfer function $t(s)$ can in general be written as the ratio of two polynomials $N(s)$ and $D(s)$:

$$t(s) = \frac{N(s)}{D(s)}$$

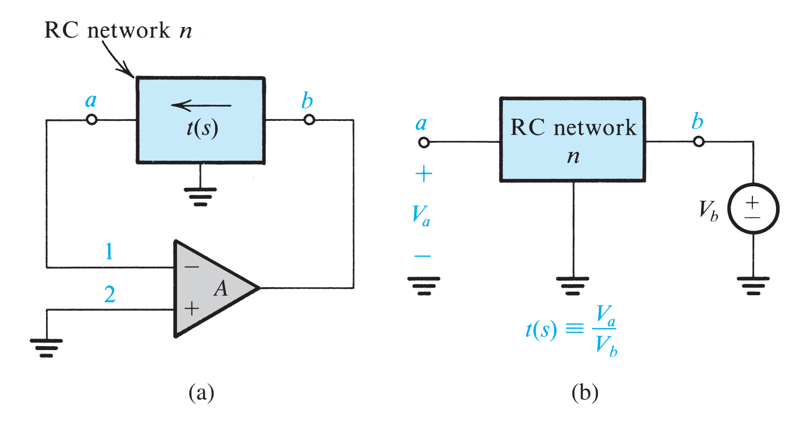


Figure 17.27 **(a)** Feedback loop obtained by placing a two-port RC network n in the feedback path of an op amp. **(b)** Definition of the open-circuit transfer function $t(s)$ of the RC network.

The roots of $N(s)$ are the transmission zeros of the RC network, and the roots of $D(s)$ are its poles. Study of circuit theory shows that while the poles of an RC network are restricted to lie on the negative real axis, the zeros can in general lie anywhere in the s plane.

The loop gain $L(s)$ of the feedback circuit in Fig. 17.27(a) can be determined using the method of Section 11.3.3. It is simply the product of the op-amp gain A and the transfer function $t(s)$,

$$L(s) = At(s) = \frac{AN(s)}{D(s)} \tag{17.69}$$

Substituting for $L(s)$ into the characteristic equation

$$1 + L(s) = 0 \tag{17.70}$$

results in the poles s_P of the closed-loop circuit obtained as solutions to the equation

$$t(s_P) = -\frac{1}{A} \tag{17.71}$$

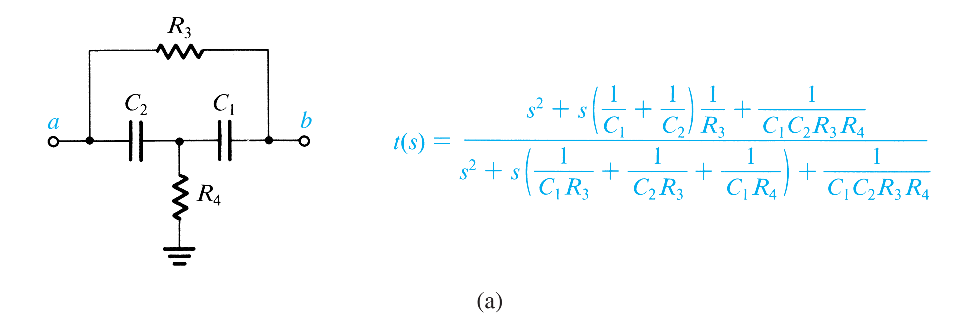
In the ideal case, $A = \infty$ and the poles are obtained from

$$N(s_P) = 0 \tag{17.72}$$

That is, *the filter poles are identical to the zeros of the RC network.*

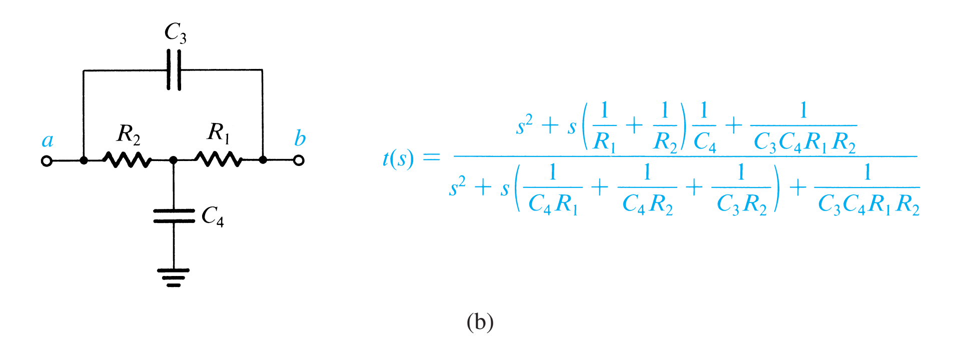
Since our objective is to realize a pair of complex-conjugate poles, we should select an RC network that can have complex-conjugate transmission zeros. The simplest such networks are the bridged-T networks shown in Fig. 17.28 together with their transfer functions $t(s)$ from b to a, with a open-circuited. As an example, consider the circuit generated by placing the bridged-T network of Fig. 17.28(a) in the negative-feedback path of an op amp, as shown in Fig. 17.29. The pole polynomial of the active-filter circuit will be equal to the numerator polynomial of the bridged-T network; thus,

$$s^2 + s\frac{\omega_0}{Q} + \omega_0^2 = s^2 + s\left(\frac{1}{C_1} + \frac{1}{C_2}\right)\frac{1}{R_3} + \frac{1}{C_1C_2R_3R_4}$$



$$t(s) = \frac{s^2 + s\left(\frac{1}{C_1} + \frac{1}{C_2}\right)\frac{1}{R_3} + \frac{1}{C_1 C_2 R_3 R_4}}{s^2 + s\left(\frac{1}{C_1 R_3} + \frac{1}{C_2 R_3} + \frac{1}{C_1 R_4}\right) + \frac{1}{C_1 C_2 R_3 R_4}}$$

(a)



$$t(s) = \frac{s^2 + s\left(\frac{1}{R_1} + \frac{1}{R_2}\right)\frac{1}{C_4} + \frac{1}{C_3 C_4 R_1 R_2}}{s^2 + s\left(\frac{1}{C_4 R_1} + \frac{1}{C_4 R_2} + \frac{1}{C_3 R_2}\right) + \frac{1}{C_3 C_4 R_1 R_2}}$$

(b)

Figure 17.28 Two RC networks (called bridged-T networks) that can have complex transmission zeros. The transfer functions given are from *b* to *a*, with *a* open-circuited.

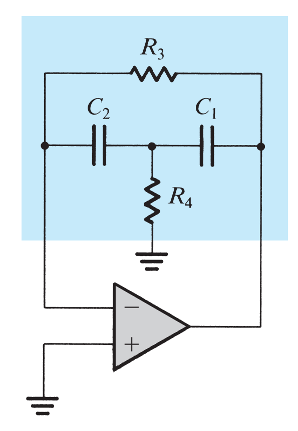


Figure 17.29 An active-filter feedback loop generated using the bridged-T network of Fig. 17.28(a).

which enables us to obtain ω_0 and Q as

$$\omega_0 = \frac{1}{\sqrt{C_1 C_2 R_3 R_4}} \tag{17.73}$$

$$Q = \left[\frac{\sqrt{C_1 C_2 R_3 R_4}}{R_3}\left(\frac{1}{C_1} + \frac{1}{C_2}\right)\right]^{-1} \tag{17.74}$$

If we are designing this circuit, ω_0 and Q are given and Eqs. (17.73) and (17.74) can be used to determine C_1, C_2, R_3, and R_4. It follows that there are two degrees of freedom. Let us

exhaust one of these by selecting $C_1 = C_2 = C$. Let us also denote $R_3 = R$ and $R_4 = R/m$. By substituting in Eqs. (17.73) and (17.74), and with some manipulation, we obtain

$$m = 4Q^2 \tag{17.75}$$

$$CR = \frac{2Q}{\omega_0} \tag{17.76}$$

Thus if we are given the value of Q, Eq. (17.75) can be used to determine the ratio of the two resistances R_3 and R_4. Then the given values of ω_0 and Q can be substituted in Eq. (17.76) to determine the time constant CR. There remains one degree of freedom—the value of C or R can be arbitrarily chosen. In an actual design, this value, which sets the *impedance level* of the circuit, should be chosen so that the resulting component values are practical.

EXERCISES

D17.25 Design the circuit of Fig. 17.29 to realize a pair of poles with $\omega_0 = 10^4$ rad/s and $Q = 1$. Select $C_1 = C_2 = 1$ nF.
Ans. $R_3 = 200$ kΩ; $R_4 = 50$ kΩ

17.26 For the circuit designed in Exercise 17.25, find the location of the poles of the RC network in the feedback loop.
Ans. -0.382×10^4 and -2.618×10^4 rad/s

17.8.2 Injecting the Input Signal

Having synthesized a feedback loop that realizes a given pair of poles, we now consider connecting the input signal source to the circuit. We wish to do this, of course, without altering the poles.

Since, for the purpose of finding the poles of a circuit, an ideal voltage source is equivalent to a short circuit, it follows that any circuit node that is connected to ground can instead be connected to the input voltage source without causing the poles to change. Thus the method of injecting the input voltage signal into the feedback loop is simply to disconnect a component (or several components) that is (are) connected to ground and connect it (them) to the input source. Depending on the component(s) through which the input signal is injected, different transmission zeros are obtained. This is, of course, the same method we used in Section 17.5 with the LCR resonator and in Section 17.6 with the biquads based on the LCR resonator.

As an example, consider the feedback loop of Fig. 17.29. Here we have two grounded nodes (one terminal of R_4 and the positive input terminal of the op amp) that can serve for injecting the input signal. Figure 17.30(a) shows the circuit with the input signal injected through part of the resistance R_4. Note that the two resistances R_4/α and $R_4/(1-\alpha)$ have a parallel equivalent of R_4.

Analysis of the circuit to determine its voltage-transfer function $T(s) \equiv V_o(s)/V_i(s)$ is illustrated in Fig. 17.30(b). Note that we have assumed the op amp to be ideal and have

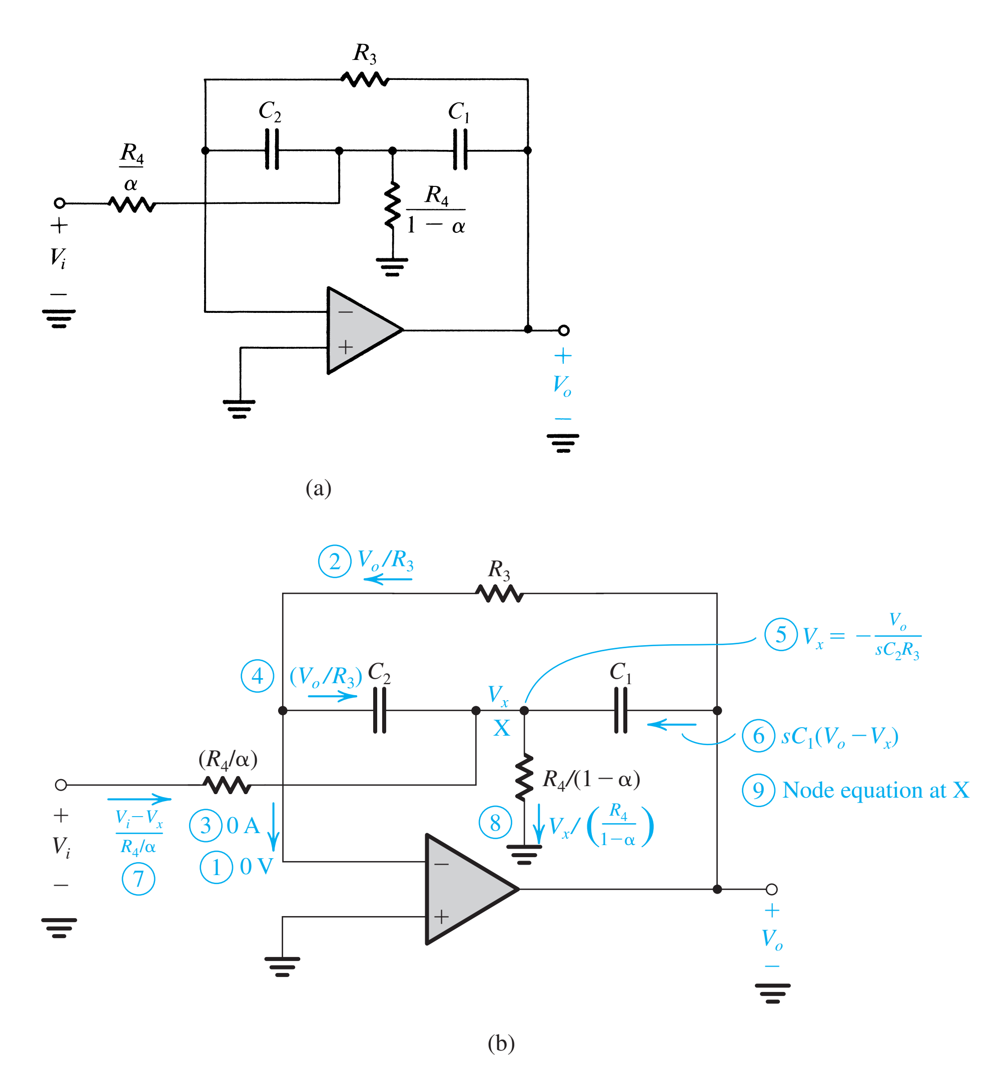


Figure 17.30 **(a)** The feedback loop of Fig. 17.29 with the input signal injected through part of resistance R_4. This circuit realizes the bandpass function. **(b)** Analysis of the circuit in **(a)** to determine its voltage transfer function $T(s)$ with the order of the analysis steps indicated by the circled numbers.

indicated the order of the analysis steps by the circled numbers. The final step, number 9, consists of writing a node equation at X and substituting for V_x by the value determined in step 5. The result is the transfer function

$$\frac{V_o}{V_i} = \frac{-s(\alpha/C_1R_4)}{s^2 + s\left(\frac{1}{C_1} + \frac{1}{C_2}\right)\frac{1}{R_3} + \frac{1}{C_1C_2R_3R_4}}$$

We recognize this as a bandpass function whose center-frequency gain can be controlled by the value of α. As expected, the denominator polynomial is identical to the numerator polynomial of $t(s)$ given in Fig. 17.28(a).

EXERCISE

17.27 Use the component values obtained in Exercise 17.25 to design the bandpass circuit of Fig. 17.30(a). Determine the values of (R_4/α) and $R_4/(1-\alpha)$ to obtain a center-frequency gain of unity.
Ans. 100 kΩ; 100 kΩ

17.8.3 Generation of Equivalent Feedback Loops

The **complementary transformation** of feedback loops is based on the property of linear networks illustrated in Fig. 17.31 for the two-port (three-terminal) network n. In Fig. 17.31(a), terminal c is grounded and a signal V_b is applied to terminal b. The transfer function from b to a with c grounded is denoted t. Then, in Fig. 17.31(b), terminal b is grounded and the input signal is applied to terminal c. The transfer function from c to a with b grounded can be shown to be the complement of t—that is, $1-t$. (Recall that we used this property in generating a circuit realization for the all-pass function in Section 17.6.)

Application of the complementary transformation to a feedback loop to generate an equivalent feedback loop is a two-step process:

1. Nodes of the feedback network and any of the op-amp inputs that are connected to ground should be disconnected from ground and connected to the op-amp output. Conversely, those nodes that were connected to the op-amp output should be now connected to ground. That is, we simply interchange the op-amp output terminal with ground.
2. The two input terminals of the op amp should be interchanged.

The feedback loop generated by this transformation has the same characteristic equation, and hence the same poles, as the original loop.

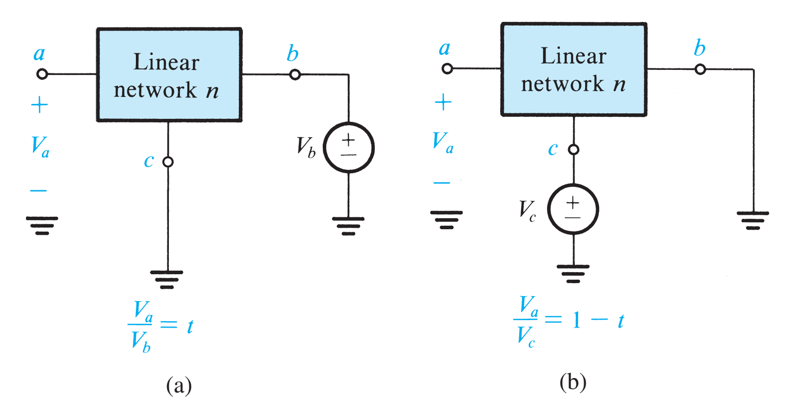
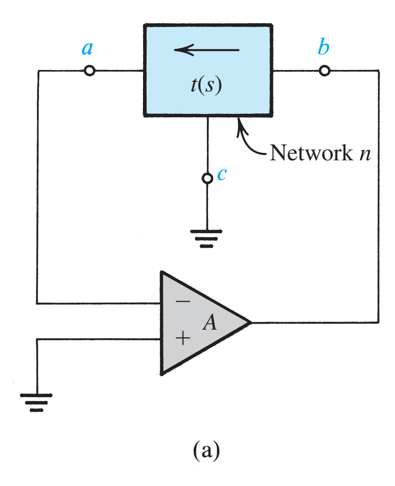


Figure 17.31 Interchanging input and ground results in the complement of the transfer function.



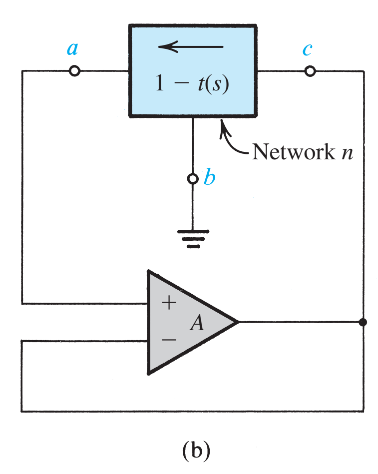


Figure 17.32 Application of the complementary transformation to the feedback loop in **(a)** results in the equivalent loop (same poles) shown in **(b)**.

To illustrate, we show in Fig. 17.32(a) the feedback loop formed by connecting a two-port RC network in the negative-feedback path of an op amp. Application of the complementary transformation to this loop results in the feedback loop of Fig. 17.32(b). Note that in the latter loop the op amp is used in the unity-gain follower configuration. We shall now show that the two loops of Fig. 17.32 are equivalent.

If the op amp has an open-loop gain A, the follower in the circuit of Fig. 17.32(b) will have a gain of $A/(A+1)$. This, together with the fact that the transfer function of network n from c to a is $1 - t$ (see Fig. 17.31), enables us to write for the circuit in Fig. 17.32(b) the characteristic equation

$$1 - \frac{A}{A+1}(1-t) = 0$$

This equation can be manipulated to the form

$$1 + At = 0$$

which is the characteristic equation of the loop in Fig. 17.32(a). As an example, consider the application of the complementary transformation to the feedback loop of Fig. 17.29: The feedback loop of Fig. 17.33(a) results. Injecting the input signal through C_1 results in the circuit in Fig. 17.33(b), which can be shown (by direct analysis) to realize a second-order high-pass function. This circuit is one of a family of SABs known as the **Sallen-and-Key circuits**, after their originators. The design of the circuit in Fig. 17.33(b) is based on Eqs. (17.73) through (17.76): namely, $R_3 = R$, $R_4 = R/4Q^2$, $C_1 = C_2 = C$, $CR = 2Q/\omega_0$, and the value of C is arbitrarily chosen to be practically convenient.

As another example, Fig. 17.34(a) shows the feedback loop generated by placing the two-port RC network of Fig. 17.28(b) in the negative-feedback path of an op amp. For an ideal op amp, this feedback loop realizes a pair of complex-conjugate natural modes having

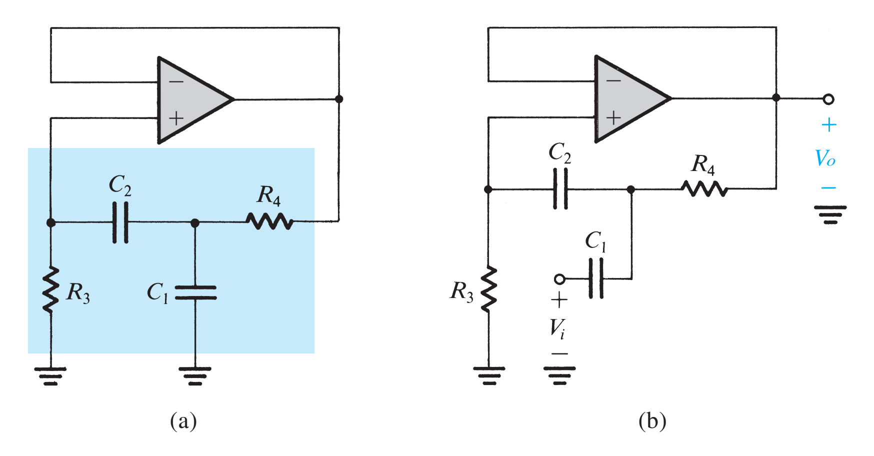


Figure 17.33 **(a)** Feedback loop obtained by applying the complementary transformation to the loop in Fig. 17.29. **(b)** Injecting the input signal through C_1 realizes the high-pass function. This is one of the Sallen-and-Key family of circuits.

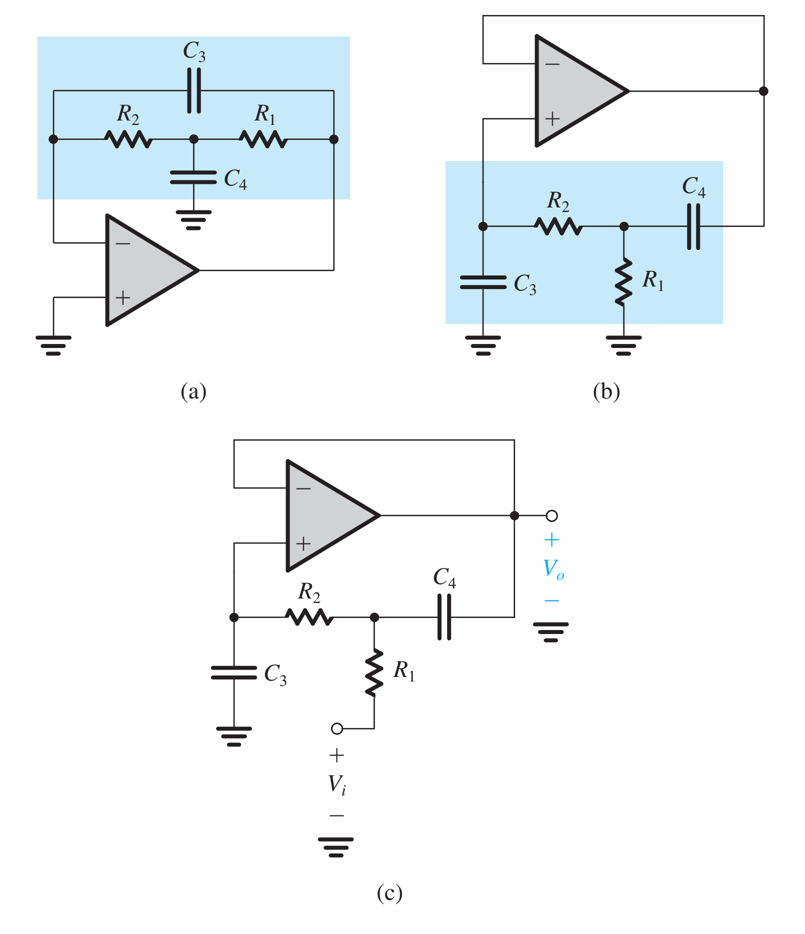


Figure 17.34 **(a)** Feedback loop obtained by placing the bridged-T network of Fig. 17.28(b) in the negative-feedback path of an op amp. **(b)** Equivalent feedback loop generated by applying the complementary transformation to the loop in **(a)**. **(c)** A low-pass filter obtained by injecting V_i through R_1 into the loop in **(b)**.

the same location as the zeros of $t(s)$ of the RC network. Thus, using the expression for $t(s)$ given in Fig. 17.28(b), we can write for the active-filter poles

$$\omega_0 = 1/\sqrt{C_3 C_4 R_1 R_2} \tag{17.77}$$

$$Q = \left[\frac{\sqrt{C_3 C_4 R_1 R_2}}{C_4}\left(\frac{1}{R_1}+\frac{1}{R_2}\right)\right]^{-1} \tag{17.78}$$

Normally the design of this circuit is based on selecting $R_1 = R_2 = R$, $C_4 = C$, and $C_3 = C/m$. When substituted in Eqs. (17.77) and (17.78), these yield

$$m = 4Q^2 \tag{17.79}$$

$$CR = 2Q/\omega_0 \tag{17.80}$$

with the remaining degree of freedom (the value of C or R) left to the designer to choose.

Injecting the input signal to the C_4 terminal that is connected to ground can be shown to result in a bandpass realization. If, however, we apply the complementary transformation to the feedback loop in Fig. 17.34(a), we obtain the equivalent loop in Fig. 17.34(b). The loop equivalence means that the circuit of Fig. 17.34(b) has the same poles and thus the same ω_0 and Q and the same design equations (Eqs. 17.77 through 17.80). The new loop in Fig. 17.34(b) can be used to realize a low-pass function by injecting the input signal as shown in Fig. 17.34(c).

In conclusion, we note that complementary transformation is a powerful tool that enables us to obtain new filter circuits from ones we already have, thus increasing our repertoire of filter realizations.

EXERCISES

17.28 Analyze the circuit in Fig. 17.34(c) to determine its transfer function $V_o(s)/V_i(s)$ and thus show that ω_0 and Q are indeed those in Eqs. (17.77) and (17.78). Also show that the dc gain is unity.

D17.29 Design the circuit in Fig. 17.34(c) to realize a low-pass filter with $f_0 = 4$ kHz and $Q = 1/\sqrt{2}$. Use 10-kΩ resistors.

Ans. $R_1 = R_2 = 10$ kΩ; $C_3 = 2.81$ nF; $C_4 = 5.63$ nF

17.9 Sensitivity

Because of the tolerances in component values and because of the finite op-amp gain, the response of the actual assembled filter will deviate from the ideal response. As a means for predicting such deviations, the filter designer employs the concept of **sensitivity**. Specifically, for second-order filters one is usually interested in finding how *sensitive* their poles are relative to variations (both initial tolerances and future drifts) in RC component values and amplifier gain. These sensitivities can be quantified using the **classical sensitivity function** S_x^y, defined as

$$S_x^y \equiv \lim_{\Delta x \to 0} \frac{\Delta y/y}{\Delta x/x} \tag{17.81}$$

Thus,

$$S_x^y = \frac{\partial y}{\partial x}\frac{x}{y} \tag{17.82}$$

Here, x denotes the value of a component (a resistor, a capacitor, or an amplifier gain) and y denotes a circuit parameter of interest (say, ω_0 or Q). For small changes

$$S_x^y \simeq \frac{\Delta y/y}{\Delta x/x} \tag{17.83}$$

Thus we can use the value of S_x^y to determine the per-unit change in y due to a given per-unit change in x. For instance, if the sensitivity of Q relative to a particular resistance R_1 is 5, then a 1% increase in R_1 results in a 5% increase in the value of Q.

Example 17.3

For the feedback loop of Fig. 17.29, find the sensitivities of ω_0 and Q relative to all the passive components and the op-amp gain. Evaluate these sensitivities for the design considered in the preceding section for which $C_1 = C_2$.

Solution

To find the sensitivities with respect to the passive components, called **passive sensitivities**, we assume that the op-amp gain is infinite. In this case, ω_0 and Q are given by Eqs. (17.73) and (17.74). Thus for ω_0 we have

$$\omega_0 = \frac{1}{\sqrt{C_1C_2R_3R_4}}$$

which can be used together with the sensitivity definition of Eq. (17.82) to obtain

$$S_{C_1}^{\omega_0} = S_{C_2}^{\omega_0} = S_{R_3}^{\omega_0} = S_{R_4}^{\omega_0} = -\frac{1}{2}$$

For Q we have

$$Q = \left[\sqrt{C_1C_2R_3R_4}\left(\frac{1}{C_1} + \frac{1}{C_2}\right)\frac{1}{R_3}\right]^{-1}$$

to which we apply the sensitivity definition to obtain

$$S_{C_1}^{Q} = \frac{1}{2}\left(\sqrt{\frac{C_2}{C_1}} - \sqrt{\frac{C_1}{C_2}}\right)\left(\sqrt{\frac{C_2}{C_1}} + \sqrt{\frac{C_1}{C_2}}\right)^{-1}$$

For the design with $C_1 = C_2$ we see that $S_{C_1}^{Q} = 0$. Similarly, we can show that

$$S_{C_2}^{Q} = 0, \quad S_{R_3}^{Q} = \frac{1}{2}, \quad S_{R_4}^{Q} = -\frac{1}{2}$$

It is important to remember that the sensitivity expression should be derived *before* values corresponding to a particular design are substituted.

Example 17.3 *continued*

Next we consider the sensitivities relative to the amplifier gain. If we assume the op amp to have a finite gain A, the characteristic equation for the loop becomes

$$1 + At(s) = 0 \tag{17.84}$$

where $t(s)$ is given in Fig. 17.28(a). To simplify matters we can substitute for the passive components by their design values. This causes no errors in evaluating sensitivities, since we are now finding the sensitivity with respect to the amplifier gain. Using the design values obtained earlier—namely, $C_1 = C_2 = C$, $R_3 = R$, $R_4 = R/4Q^2$, and $CR = 2Q/\omega_0$—we get

$$t(s) = \frac{s^2 + s(\omega_0/Q) + \omega_0^2}{s^2 + s(\omega_0/Q)(2Q^2+1) + \omega_0^2} \tag{17.85}$$

where ω_0 and Q denote the nominal or design values of the pole frequency and Q factor. The actual values are obtained by substituting for $t(s)$ in Eq. (17.84):

$$s^2 + s\frac{\omega_0}{Q}\left(2Q^2+1\right) + \omega_0^2 + A\left(s^2 + s\frac{\omega_0}{Q} + \omega_0^2\right) = 0$$

Assuming the gain A to be real and dividing both sides by $A+1$, we get

$$s^2 + s\frac{\omega_0}{Q}\left(1 + \frac{2Q^2}{A+1}\right) + \omega_0^2 = 0 \tag{17.86}$$

From this equation we see that the actual pole frequency, ω_{0a}, and the pole Q, Q_a, are

$$\omega_{0a} = \omega_0 \tag{17.87}$$

$$Q_a = \frac{Q}{1 + 2Q^2/(A+1)} \tag{17.88}$$

Thus

$$S_A^{\omega_{0a}} = 0$$

$$S_A^{Q_a} = \frac{A}{A+1}\,\frac{2Q^2/(A+1)}{1 + 2Q^2/(A+1)}$$

For $A \gg 2Q^2$ and $A \gg 1$ we obtain

$$S_A^{Q_a} \simeq \frac{2Q^2}{A}$$

It is usual to drop the subscript a in this expression and write

$$S_A^{Q} \simeq \frac{2Q^2}{A} \tag{17.89}$$

Note that if Q is high ($Q \geq 5$), its sensitivity relative to the amplifier gain can be quite high.[10]

---

[10] Because the open-loop gain A of op amps usually has wide tolerance, it is important to keep $S_A^{\omega_0}$ and S_A^{Q} very small.

The results of Example 17.3 indicate a serious disadvantage of single-amplifier biquads—the sensitivity of Q relative to the amplifier gain is quite high. Although a technique exists for reducing S_A^Q in SABs (see Sedra et al., 1980), this is done at the expense of increased passive sensitivities. Nevertheless, the resulting SABs are used extensively in many applications. However, for filters with Q factors greater than about 10, one usually opts for one of the multiamplifier biquads studied in Sections 17.6 and 17.7. For these circuits S_A^Q is proportional to Q, rather than to Q^2 as in the SAB case (Eq. 17.89).

EXERCISE

17.30 In a particular filter utilizing the feedback loop of Fig. 17.29, with $C_1 = C_2$, use the results of Example 17.3 to find the expected percentage change in ω_0 and Q under the conditions that (a) R_3 is 2% high, (b) R_4 is 2% high, (c) both R_3 and R_4 are 2% high, and (d) both capacitors are 2% low and both resistors are 2% high.
Ans. (a) –1%, +1%; (b) –1%, –1%; (c) –2%, 0%; (d) 0%, 0%

17.10 Transconductance-C Filters

The op amp–RC circuits studied in Sections 17.4 and 17.6 through 17.8 are ideally suited for implementing audio-frequency filters using discrete op amps, resistors, and capacitors, assembled on printed-circuit boards. Such circuits have also been implemented in hybrid thin- or thick-film forms where the op amps are used in chip form (i.e., without their packages).

The limitation of op amp–RC filters to low-frequency applications is a result of the relatively low bandwidth of general-purpose op amps. The lack of suitability of these filter circuits for implementation in IC form stems from:

1. The need for large-valued capacitors, which would require impractically large chip areas;
2. The need for very precise values of RC time constants. This is impossible to achieve on an IC without resorting to expensive trimming and tuning techniques; and
3. The need for op amps that can drive resistive and large capacitive loads. As we have seen, CMOS op amps are usually capable of driving only small capacitances.

17.10.1 Methods for IC Filter Implementation

We now introduce the three approaches currently in use for implementing filters in monolithic form.

Transconductance-C Filters These utilize transconductance amplifiers or simply transconductors together with capacitors and are hence called G_m–C filters. Because high-quality and high-frequency transconductors can be easily realized in CMOS technology, where small-valued capacitors are plentiful, this filter-design method is very popular at this time. It

has been used at medium and high frequencies approaching the hundreds of megahertz range. We shall study this method briefly in this section.

MOSFET-C Filters These utilize the two-integrator-loop circuits of Section 17.8 but with the resistors replaced with MOSFETs operating in the triode region. Clever techniques have been evolved to obtain linear operation with large input signals. Because of space limitations, we shall not study this design method here and refer the reader to Tsividis and Voorman (1992).

Switched-Capacitor Filters These are based on the ingenious technique of obtaining a large resistance by switching a capacitor at a relatively high frequency. Because of the switching action, the resulting filters are discrete-time circuits, as opposed to the continuous-time filters studied thus far. The switched-capacitor approach is ideally suited for implementing low-frequency filters in IC form using CMOS technology. We shall study switched-capacitor filters in Section 17.11.

EARLY FILTER PIONEERS—CAUER AND DARLINGTON:

While on a fellowship with Vannevar Bush at MIT and Harvard, the German mathematician Wilhelm Cauer (1900–1945) used the Chebyshev polynomials in a way that unified the field of filter transfer function design. The elliptical filters now known as Cauer filters have equiripple performance in both the passband and the stopband(s). Cauer continued to make contributions to LC filter synthesis until his mysterious disappearance and presumed death in Berlin on the last day of the Second World War.

Sidney Darlington (1906–1997) developed a complete design theory for LC filters while working at the Bell Telephone Laboratories in the 1940s. Ironically, in later years he became better known for his invention of a particular transistor circuit, the Darlington pair.

17.10.2 Transconductors

Figure 17.35(a) shows the circuit symbol for a transconductor, and Fig. 17.35(b) shows its equivalent circuit. Here we are assuming the transconductor to be ideal, with infinite input and output impedances. Actual transconductors will obviously deviate from this ideal model. We shall investigate the effects of nonidealities in some of the end-of-chapter problems. Otherwise, we shall assume that for the purpose of this introductory study, the transconductors are ideal.

The transconductor of Fig. 17.35(a) has a positive output; that is, the output current $I_o = G_m V_i$ flows *out* of the output terminal. Transconductors with a negative output are, of course, also possible and one is shown in Fig. 17.35(c), with its ideal model in Fig. 17.35(d).

The transconductors of Fig. 17.35(a) and (c) are both of the single-ended type. As mentioned in Chapter 9, differential amplification is preferred over the single-ended variety for a number of reasons, including lower susceptibility to noise and interference. This preference for fully differential operation extends to other signal-processing functions including filtering where it can be shown that distortion, an important issue in filter design, is reduced in fully differential configurations. As a result, at the present time, most IC analog filters utilize fully differential circuits. For this purpose, we show in Fig. 17.35(e) a differential-input–differential-output transconductor.

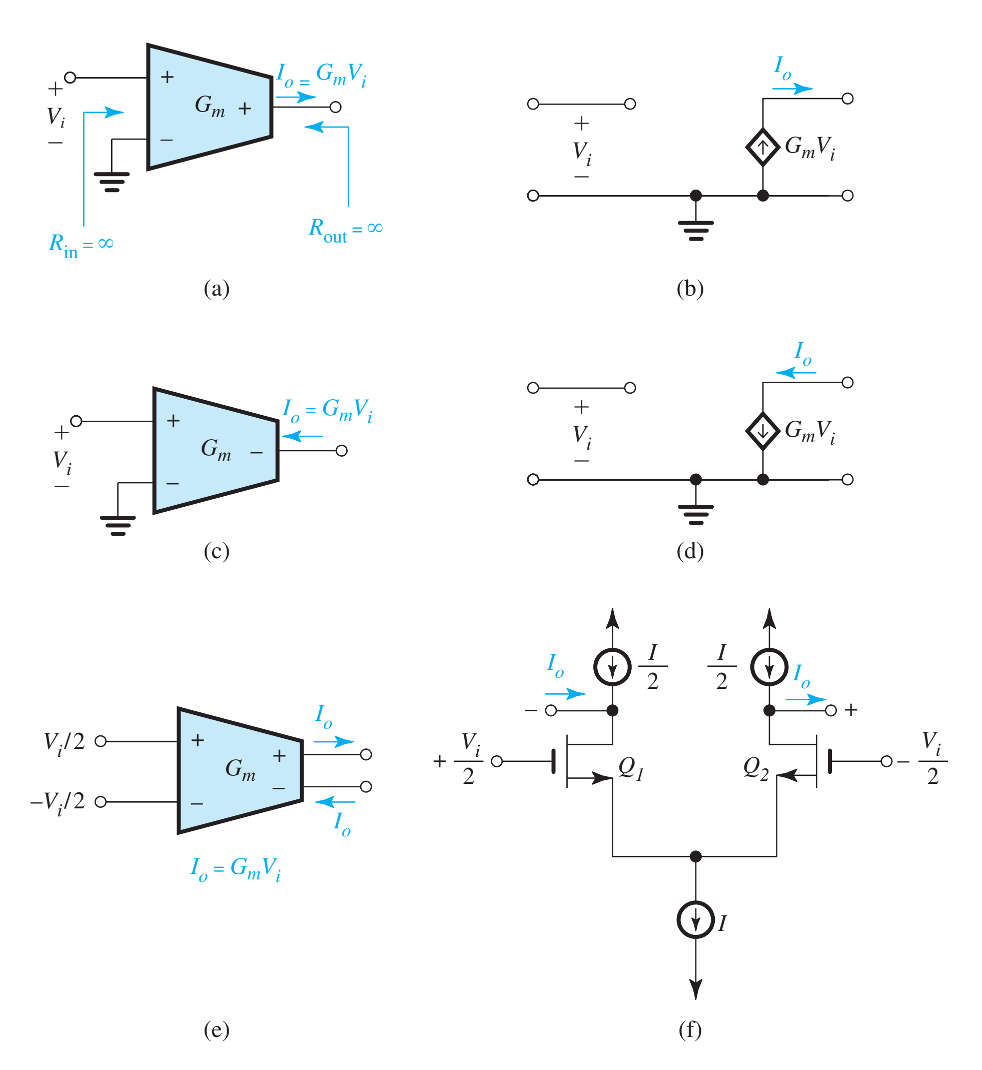


Figure 17.35 (**a**) A positive transconductor; (**b**) equivalent circuit of the transconductor in (**a**); (**c**) a negative transconductor and its equivalent circuit (**d**); (**e**) a fully differential transconductor; (**f**) a simple circuit implementation of the fully differential transconductor.

We have already encountered circuits for implementing transconductors. As an example of a simple implementation, we show the circuit in Fig. 17.35(f), which is simply a differential amplifier loaded with two current sources. The linearity of this circuit is of course limited by the i_D–v_{GS} characteristic of the MOSFET, necessitating the use of small input signals. Many elaborate transconductor circuits have been proposed and utilized in the design of G_m–C filters (see Chan Carusone, Johns, and Martin, 2012).

17.10.3 Basic Building Blocks

In this section we present the basic building blocks of G_m–C filters. Figure 17.36(a) shows how a negative transconductor can be used to realize a resistance. An integrator is obtained by feeding the output current of a transconductor, $G_m V_i$, to a grounded capacitor, as shown in

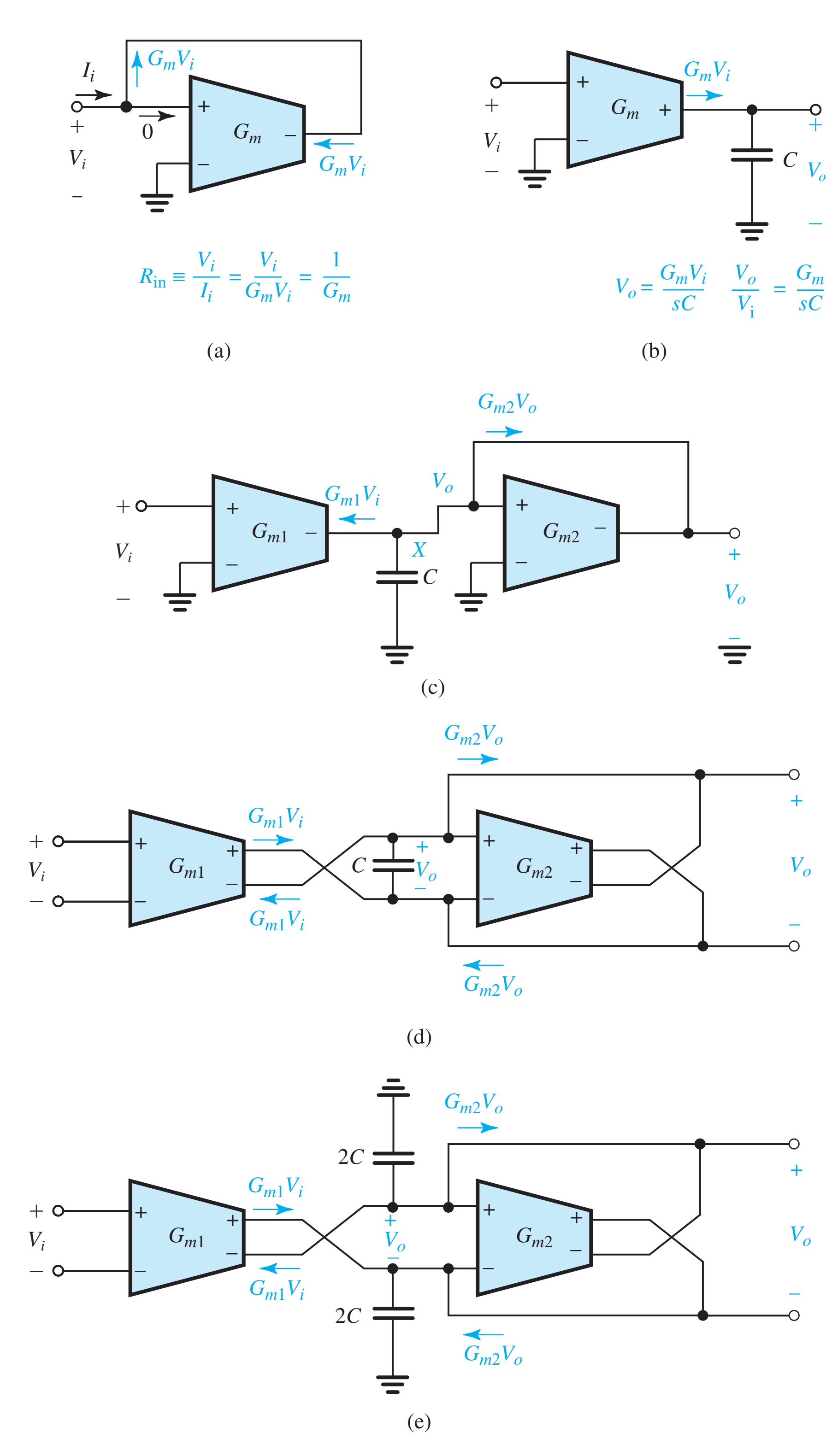


Figure 17.36 Realization of (**a**) a resistance using a negative transconductor; (**b**) an ideal noninverting integrator; (**c**) a first-order low-pass filter (a damped integrator); and (**d**) a fully differential first-order low-pass filter. (**e**) Alternative realization of the fully differential first-order low-pass filter.

Fig. 17.36(b). The transfer function obtained is

$$\frac{V_o}{V_i} = \frac{G_m}{sC} \tag{17.90}$$

which is ideal because we have assumed the transconductor to be ideal (see Problem 17.86).

To obtain a damped integrator, or a first-order low-pass filter, we connect a resistance of the type in Fig. 17.36(a) in parallel with the capacitor C in the integrator of Fig. 17.36(b). The resulting circuit is shown in Fig. 17.36(c). The transfer function can be obtained by writing a node equation at X. The result is

$$\frac{V_o}{V_i} = -\frac{G_{m1}}{sC + G_{m2}} \tag{17.91}$$

Thus, the pole frequency is (G_{m2}/C) and the dc gain is $(-G_{m1}/G_{m2})$.

The circuit in Figure 17.36(c) can be easily converted to the fully differential form shown in Fig. 17.36(d). An alternative implementation of the fully differential first-order low-pass filter is shown in Fig. 17.36(e). Note that the latter circuit requires four times the capacitance value of the circuit in Fig. 17.36(d). Nevertheless, the circuit of Fig. 17.36(e) has some advantages (see Chan Carusone et al., 2012).

17.10.4 Second-Order G_m–C Filter

To obtain a second-order G_m–C filter, we use the two-integrator-loop topology of Fig. 17.25(a). Absorbing the $(1/Q)$ branch within the first integrator, and lumping the second integrator together with the inverter into a single noninverting integrator block, we obtain the block diagram in Fig. 17.37(a). This block diagram can be easily implemented by G_m–C circuits, resulting in the circuit of Fig. 17.37(b). Note that

1. The inverting integrator is realized by the inverting transconductor G_{m1}, capacitor C_1, and the resistance implemented by transconductor G_{m3}.
2. The noninverting integrator is realized by the noninverting transconductor G_{m2} and capacitor C_2.
3. The input summer is implemented by transconductor G_{m4}, which feeds an output current $G_{m4}V_i$ to the integrator capacitor C_1, and transconductor G_{m1}, which feeds an output current $G_{m1}V_2$ to C_1.

To derive the transfer functions (V_1/V_i) and (V_2/V_i) we first note that V_2 and V_1 are related by

$$V_2 = \frac{G_{m2}}{sC_2}V_1 \tag{17.92}$$

Next, we write a node equation at X and use the relationship above to eliminate V_2. After some simple algebraic manipulations we obtain

$$\frac{V_1}{V_i} = -\frac{s(G_{m4}/C_1)}{s^2 + s\dfrac{G_{m3}}{C_1} + \dfrac{G_{m1}G_{m2}}{C_1C_2}} \tag{17.93}$$

Now, using Eq. (17.92) to replace V_1 in Eq. (17.93) results in

$$\frac{V_2}{V_i} = -\frac{G_{m2}G_{m4}/C_1C_2}{s^2 + s\dfrac{G_{m3}}{C_1} + \dfrac{G_{m1}G_{m2}}{C_1C_2}} \tag{17.94}$$

Thus, the circuit in Fig. 17.37(b) is capable of realizing simultaneously a bandpass function (V_1/V_i) and a low-pass function (V_2/V_i). For both

$$\omega_0 = \sqrt{\frac{G_{m1}G_{m2}}{C_1C_2}} \tag{17.95}$$

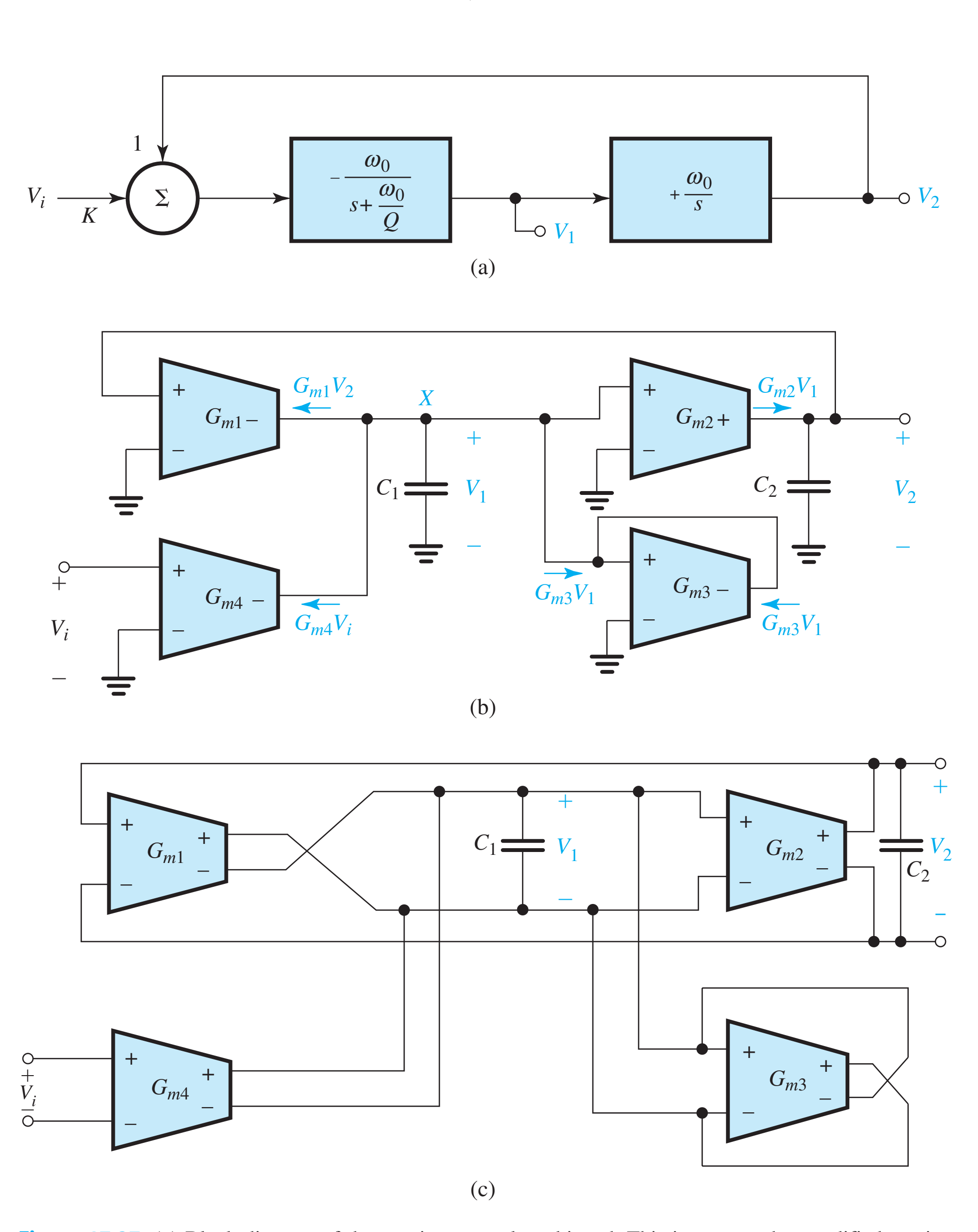


Figure 17.37 (**a**) Block diagram of the two-integrator-loop biquad. This is a somewhat modified version of Fig. 17.25. (**b**) G_m–C implementation of the block diagram in (**a**). (**c**) Fully differential G_m–C implementation of the block diagram in (**a**). In all parts, V_1/V_i is a bandpass function and V_2/V_i is a low-pass function.

and

$$Q = \frac{\sqrt{G_{m1}G_{m2}}}{G_{m3}}\sqrt{\frac{C_1}{C_2}} \tag{17.96}$$

For the bandpass function,

$$\text{Center-frequency gain} = -\frac{G_{m4}}{G_{m3}} \tag{17.97}$$

and for the low-pass function,

$$\text{DC gain} = -\frac{G_{m4}}{G_{m1}} \tag{17.98}$$

There are a variety of possible designs. The most common is to make the time constants of the integrators equal [which is the case in the block diagram of Fig. 17.37(a)]. Doing this and selecting $G_{m1} = G_{m2} = G_m$ and $C_1 = C_2 = C$ results in the following design equation

$$\frac{G_m}{C} = \omega_0 \tag{17.99}$$

$$G_{m3} = \frac{G_m}{Q} \tag{17.100}$$

$$\text{For the BP:} \quad G_{m4} = \frac{G_m}{Q}\,|\text{Gain}| \tag{17.101}$$

$$\text{For the LP:} \quad G_{m4} = G_m\,|\text{Gain}| \tag{17.102}$$

Example 17.4

Design the G_m–C circuit of Fig. 17.37(b) to realize a bandpass filter with a center frequency of 10 MHz, a 3-dB bandwidth of 1 MHz, and a center-frequency gain of 10. Use equal capacitors of 5 pF.

Solution

Using the equal-integrator-time-constants design, Eq. (17.99) yields

$$G_m = \omega_0 C = 2\pi \times 10 \times 10^6 \times 5 \times 10^{-12} = 0.314 \text{ mA/V}$$

Thus,

$$G_{m1} = G_{m2} = 0.314 \text{ mA/V}$$

To obtain G_{m3}, we first note that $Q = f_0/\text{BW} = 10/1 = 10$, and then use Eq. (17.100) to obtain

$$G_{m3} = \frac{G_m}{Q} = \frac{0.314}{10} = 0.0314 \text{ mA/V}$$

or

$$G_{m3} = 31.4 \text{ µA/V}$$

Finally, G_{m4} can be found by using Eq. (17.101) as

$$G_{m4} = \frac{G_m}{10} \times 10 = 0.314 \text{ mA/V}$$

We note that the feedforward approach utilized in Section 17.7.3 to realize different transmission zeros (as required for high-pass, notch, and all-pass functions) can be adapted to the G_m–C circuit in Fig. 17.37(b). Some of these possibilities are explored in the end-of-chapter problems. Finally, the circuit in Fig. 17.37(b) can be easily converted to the fully differential form shown in Fig. 17.37(c).

EXERCISE

D17.31 Design the circuit of Fig. 17.37(b) to realize a maximally flat low-pass filter with $f_{3dB} = 20$ MHz and a dc gain of unity. Design for equal integrator time constants, and use equal capacitors of 2 pF each.
Ans. $G_{m1} = G_{m2} = G_{m4} = 0.251$ mA/V; $G_{m3} = 0.355$ mA/V

This concludes our study of G_m–C filters. The interested reader can find considerably more material on this subject in Schaumann et al. (2010).

17.11 Switched-Capacitor Filters

In this section we study another approach to the design of analog filters for IC implementation. Switched-capacitor filters, which require only small capacitors, analog switches, and op amps that need to drive only small capacitive loads, are ideally suited for implementation in CMOS. Currently, the switched-capacitor approach is the preferred method for the design of integrated-circuit audio filters.

17.11.1 The Basic Principle

The switched-capacitor filter technique is based on the realization that a capacitor switched between two circuit nodes at a sufficiently high rate is equivalent to a resistor connecting these two nodes. To be specific, consider the active-RC integrator of Fig. 17.38(a). This is the familiar Miller integrator, which we used in the two-integrator-loop biquad in Section 17.7. In Fig. 17.38(b) we have replaced the input resistor R_1 by a grounded capacitor C_1 together with two MOS transistors acting as switches. In some circuits, more elaborate switch configurations are used, but such details are beyond our present need.

The two MOS switches in Fig. 17.38(b) are driven by a *nonoverlapping* two-phase clock. Figure 17.38(c) shows the clock waveforms. We shall assume in this introductory exposition that the clock frequency f_c ($f_c = 1/T_c$) is much higher than the frequency of the input signal v_i. Thus the variations in the input signal are negligibly small during clock phase ϕ_1, when C_1 is connected across the input signal source v_i. It follows that during ϕ_1, capacitor C_1 charges up to the voltage v_i,

$$q_{C1} = C_1 v_i$$

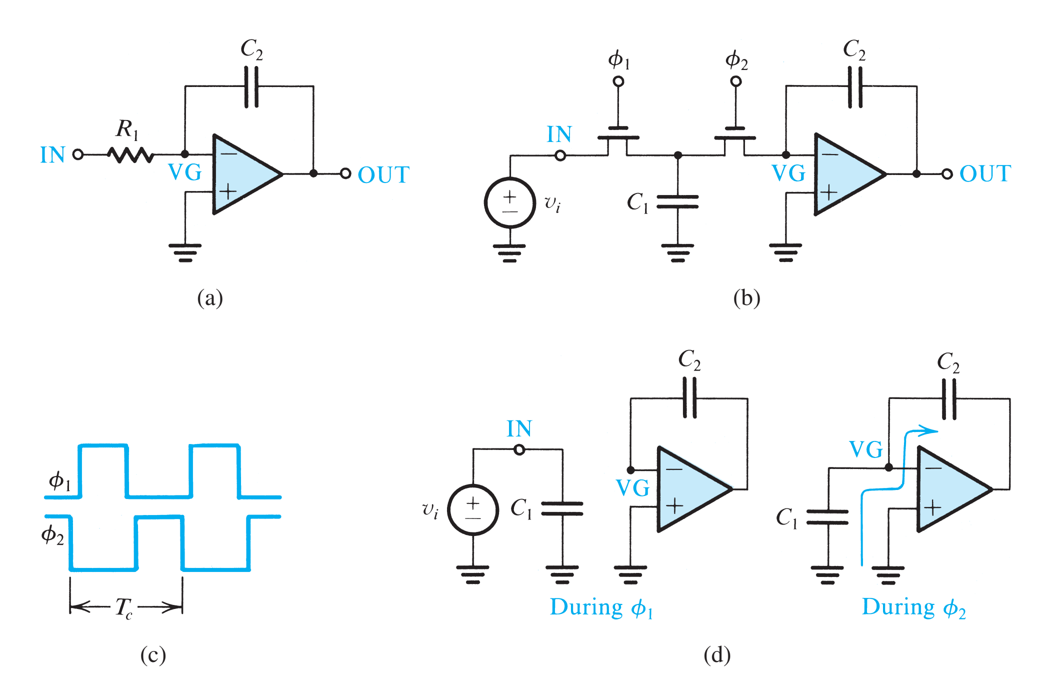


Figure 17.38 Basic principle of the switched-capacitor filter technique. **(a)** Active-RC integrator. **(b)** Switched-capacitor integrator. **(c)** Two-phase clock (nonoverlapping). **(d)** During ϕ_1, C_1 charges up to the current value of v_i and then, during ϕ_2, discharges into C_2.

Then, during clock phase ϕ_2, capacitor C_1 is connected to the virtual-ground input of the op amp, as indicated in Fig. 17.38(d). Capacitor C_1 is thus forced to discharge, and its previous charge q_{C1} is transferred to C_2, in the direction indicated in Fig. 17.38(d).

From the description above we see that during each clock period T_c an amount of charge $q_{C1} = C_1 v_i$ is extracted from the input source and supplied to the integrator capacitor C_2. Thus the average current flowing between the input node (IN) and the virtual-ground node (VG) is

$$i_{av} = \frac{C_1 v_i}{T_c}$$

If T_c is sufficiently short, one can think of this process as almost continuous and thus can define an equivalent resistance R_{eq} that is in effect present between nodes IN and VG:

$$R_{eq} \equiv v_i / i_{av}$$

Thus,

$$R_{eq} = T_c / C_1 \tag{17.103}$$

Using R_{eq} we obtain an equivalent time constant for the integrator:

$$\text{Time constant} = C_2 R_{eq} = T_c \frac{C_2}{C_1} \tag{17.104}$$

Thus the time constant that determines the frequency response of the filter is established by the clock period T_c and the capacitor ratio C_2/C_1. Both these parameters can be well controlled

in an IC process. Specifically, note the dependence on capacitor ratios rather than on absolute values of capacitors. The accuracy of capacitor ratios in MOS technology can be controlled to within 0.1%.

Another point worth observing is that with a reasonable clocking frequency (such as 100 kHz) and not-too-large capacitor ratios (say, 10), one can obtain reasonably large time constants (such as 10^{-4} s) suitable for audio applications. Since capacitors typically occupy relatively large areas on the IC chip, one attempts to minimize their values. In this context, it is important to note that the ratio accuracies quoted earlier are obtainable with the smaller capacitor value as low as 0.1 pF.

17.11.2 Practical Circuits

The switched-capacitor (SC) circuit in Fig. 17.38(b) realizes an inverting integrator [note the direction of charge flow through C_2 in Fig. 17.38(d)]. As we saw in Section 17.7, a two-integrator-loop active filter is composed of one inverting and one noninverting integrator.[11] To realize a switched-capacitor biquad filter, we therefore need a pair of complementary switched-capacitor integrators. Figure 17.39(a) shows a noninverting, or positive, integrator circuit. The reader is urged to follow the operation of this circuit during the two clock phases and thus show that it operates in much the same way as the basic circuit of Fig. 17.38(b), except for a sign reversal.

In addition to realizing a noninverting integrator function, the circuit in Fig. 17.39(a) is insensitive to stray capacitances; however, we shall not explore this point any further. The interested reader is referred to Schaumann, Ghausi, and Laker (1990). By reversal of the clock phases on two of the switches, the circuit in Fig. 17.39(b) is obtained. This circuit realizes the inverting integrator function, like the circuit of Fig. 17.38(b), but is insensitive to stray capacitances [which the original circuit of Fig. 17.38(b) is not]. The complementary integrators of Fig. 17.39 have become the standard building blocks in the design of switched-capacitor filters.

Let us now consider the realization of a complete biquad circuit. Figure 17.40(a) shows the active-RC, two-integrator-loop circuit studied earlier. By considering the cascade of integrator 2 and the inverter as a positive integrator, and then simply replacing each resistor by its switched-capacitor equivalent, we obtain the circuit in Fig. 17.40(b). Ignore the damping around the first integrator (i.e., the switched capacitor C_5) for the time being and note that the feedback loop indeed consists of one inverting and one noninverting integrator. Then note the phasing of the switched capacitor used for damping. Reversing the phases here would convert the feedback to positive and move the poles to the right half of the s plane. On the other hand, the phasing of the feed-in switched capacitor (C_6) is not that important; a reversal of phases would result only in an inversion in the sign of the function realized.

Having identified the correspondences between the active-RC biquad and the switched-capacitor biquad, we can now derive design equations. Analysis of the circuit in Fig. 17.40(a) yields

$$\omega_0 = \frac{1}{\sqrt{C_1 C_2 R_3 R_4}} \tag{17.105}$$

Replacing R_3 and R_4 with their switched-capacitor equivalent values, that is,

$$R_3 = T_c/C_3 \quad \text{and} \quad R_4 = T_c/C_4$$

[11]In the two-integrator loop of Fig. 17.25(b), the noninverting integrator is realized by the cascade of a Miller integrator and an inverting amplifier.

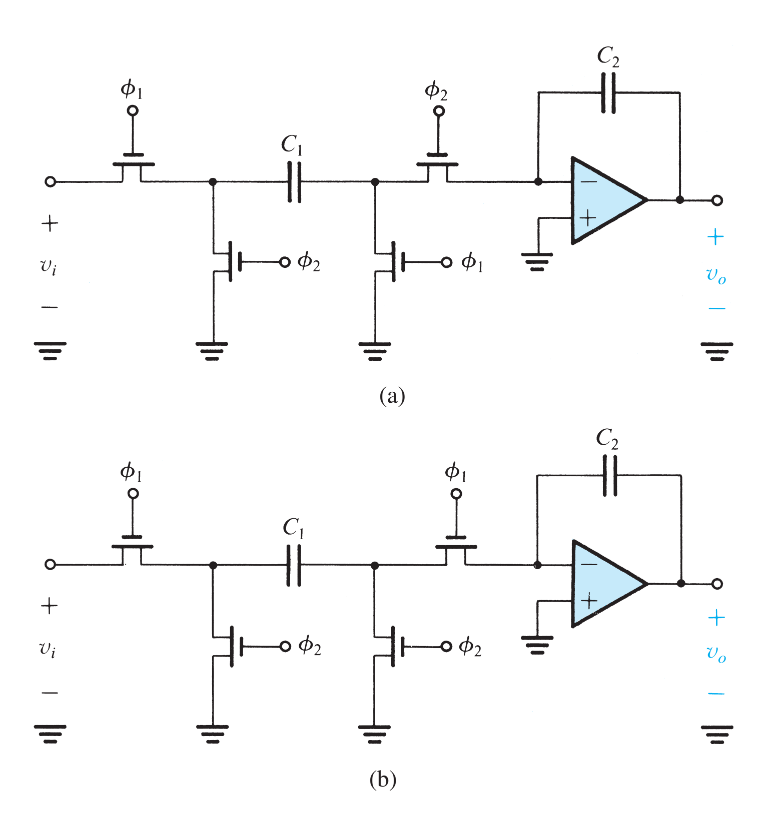


Figure 17.39 A pair of complementary stray-insensitive, switched-capacitor integrators. **(a)** Noninverting switched-capacitor integrator. **(b)** Inverting switched-capacitor integrator.

gives ω_0 of the switched-capacitor biquad as

$$\omega_0 = \frac{1}{T_c}\sqrt{\frac{C_3}{C_2}\frac{C_4}{C_1}} \tag{17.106}$$

It is usual to select the time constants of the two integrators to be equal; that is,

$$\frac{T_c}{C_3}C_2 = \frac{T_c}{C_4}C_1 \tag{17.107}$$

If, further, we select the two integrating capacitors C_1 and C_2 to be equal,

$$C_1 = C_2 = C \tag{17.108}$$

then

$$C_3 = C_4 = KC \tag{17.109}$$

where from Eq. (17.106)

$$K = \omega_0 T_c \tag{17.110}$$

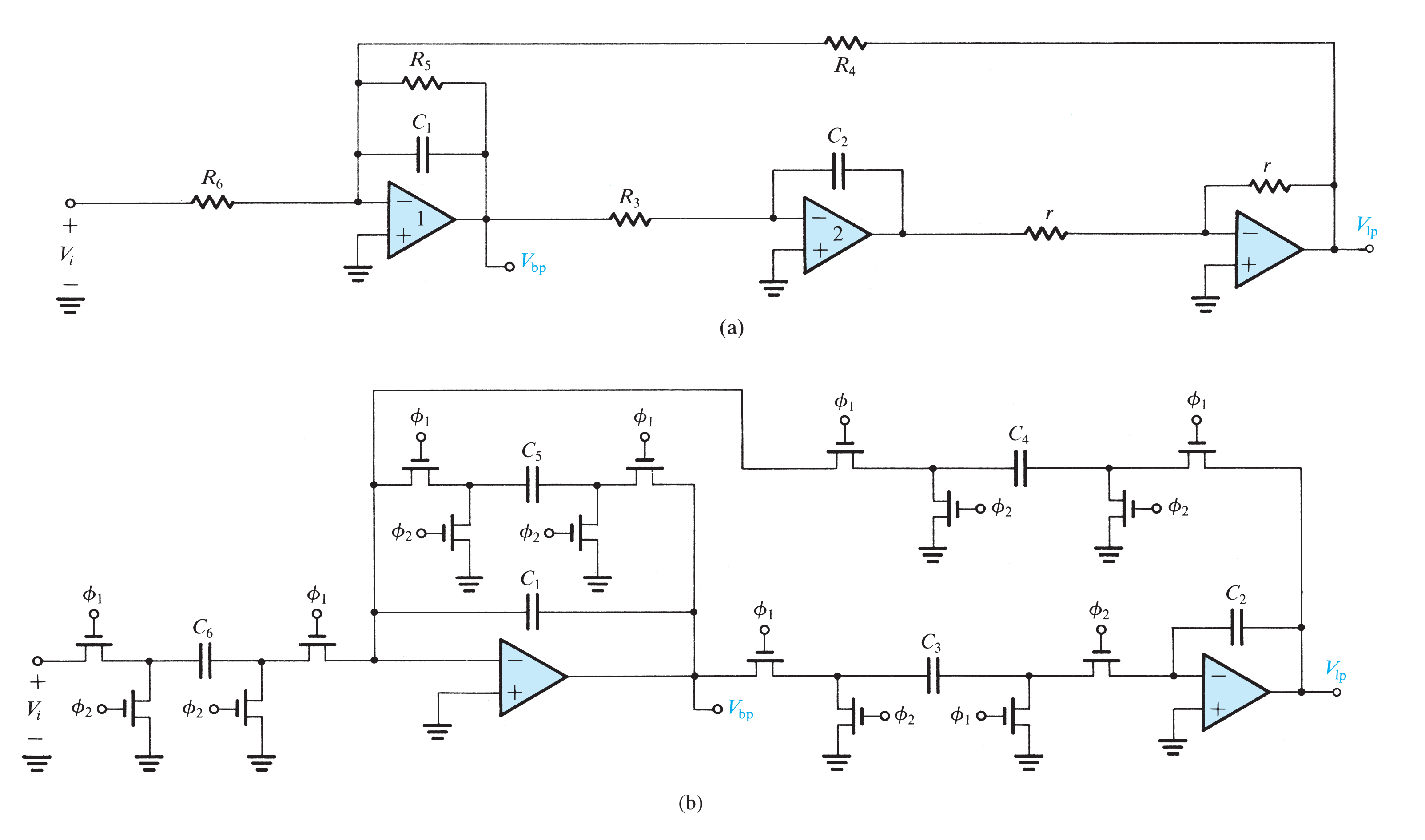


Figure 17.40 **(a)** A two-integrator-loop, active-RC biquad and **(b)** its switched-capacitor counterpart.

For the case of equal time constants, the Q factor of the circuit in Fig. 17.40(a) is given by R_5/R_4. Thus the Q factor of the corresponding switched-capacitor circuit in Fig. 17.40(b) is given by

$$Q = \frac{T_c/C_5}{T_c/C_4} = \frac{C_4}{C_5} \tag{17.111}$$

Thus C_5 should be selected from

$$C_5 = \frac{C_4}{Q} = \frac{KC}{Q} = \omega_0 T_c \frac{C}{Q} \tag{17.112}$$

Finally, the center-frequency gain of the bandpass function is given by

$$\text{Center-frequency gain} = \frac{C_6}{C_5} = Q\frac{C_6}{\omega_0 T_c C} \tag{17.113}$$

EXERCISE

D17.32 Use $C_1 = C_2 = 20$ pF and design the circuit in Fig. 17.40(b) to realize a bandpass function with $f_0 = 10$ kHz, $Q = 20$, and unity center-frequency gain. Use a clock frequency $f_c = 200$ kHz. Find the values of C_3, C_4, C_5, and C_6.
Ans. 6.283 pF; 6.283 pF; 0.314 pF; 0.314 pF

17.11.3 Final Remarks

We have attempted to provide only an introduction to switched-capacitor filters. We have made many simplifying assumptions, the most important being the switched-capacitor–resistor equivalence (Eq. 17.103). This equivalence is correct only at $f_c = \infty$ and is approximately correct for $f_c \gg f$. Switched-capacitor filters are, in fact, discrete-time circuits whose analysis and design can be carried out exactly using z-transform techniques. The interested reader is referred to the bibliography in Appendix I.

The switched-capacitor circuits presented above are of the single-ended variety. In most applications, fully differential versions of these circuits are employed.

17.12 Tuned Amplifiers

We conclude this chapter with the study of a special kind of frequency-selective network, the LC-tuned amplifier. Figure 17.41 shows the general shape of the frequency response of a tuned amplifier. The techniques discussed apply to amplifiers with center frequencies in the range of a few hundred kilohertz to a few hundred megahertz. Tuned amplifiers find application in the radio-frequency (RF) and intermediate-frequency (IF) sections of communications receivers and in a variety of other systems. It should be noted that the tuned-amplifier response of Fig. 17.41 is similar to that of the bandpass filter discussed in earlier sections.

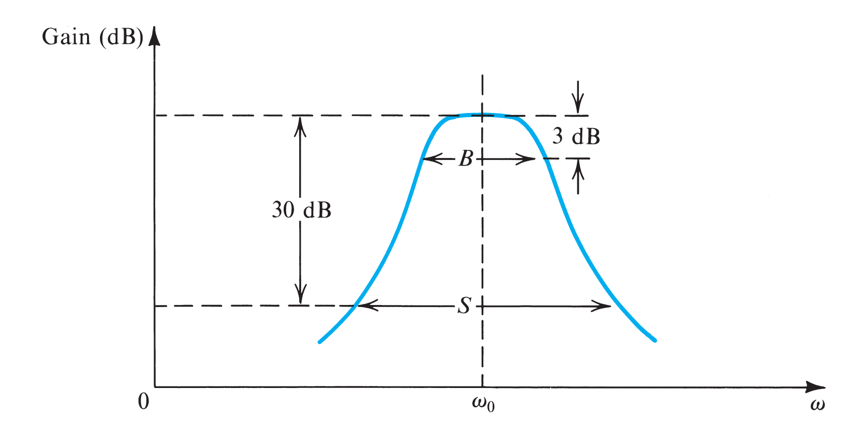


Figure 17.41 Frequency response of a tuned amplifier.

As indicated in Fig. 17.41, the response is characterized by the center frequency ω_0, the 3-dB bandwidth B, and the *skirt selectivity,* which is usually measured as the ratio of the 30-dB bandwidth to the 3-dB bandwidth. In many applications, the 3-dB bandwidth is less than 1% of ω_0. This **narrow-band** property makes possible certain approximations that can simplify the design process.

The tuned amplifiers discussed in this section can be implemented in discrete-circuit form using transistors together with passive inductors and capacitors. Increasingly, however, they are implemented in IC form, where the inductors are specially fabricated by depositing thin metal films in a spiral shape. These IC inductors, however, are very small and hence are useful only in very-high-frequency applications. Also they usually have considerable losses or, equivalently, low Q factors. Various circuit techniques have been proposed to raise the realized Q factors. These usually involve an amplifier circuit that generates a negative resistance, which is connected to the inductor in a way that cancels part of its resistance and thus enhance its Q factor. The resulting tuned amplifiers are therefore referred to as *active-LC filters* (see Schaumann et al., 2010).

This section considers tuned amplifiers that are small-signal voltage amplifiers in which the transistors operate in the "class A" mode; that is, the transistors conduct at all times. Tuned power amplifiers such as those based on class C operation of the transistor, are not studied in this book. (For a discussion on the classification of amplifiers, refer to Section 12.1.)

17.12.1 The Basic Principle

The basic principle underlying the design of tuned amplifiers is the use of a parallel LCR circuit as the load, or at the input, of a BJT or an FET amplifier. This is illustrated in Fig. 17.42 with a MOSFET amplifier having a tuned-circuit load. For simplicity, the bias details are not included. Since this circuit uses a single tuned circuit, it is known as a **single-tuned amplifier**. The amplifier equivalent circuit is shown in Fig. 17.42(b). Here R denotes the parallel equivalent of R_L and the output resistance r_o of the FET, and C is the parallel equivalent of C_L and the FET output capacitance (usually small). From the equivalent circuit we can write

$$V_o = \frac{-g_m V_i}{Y_L} = \frac{-g_m V_i}{sC + 1/R + 1/sL}$$

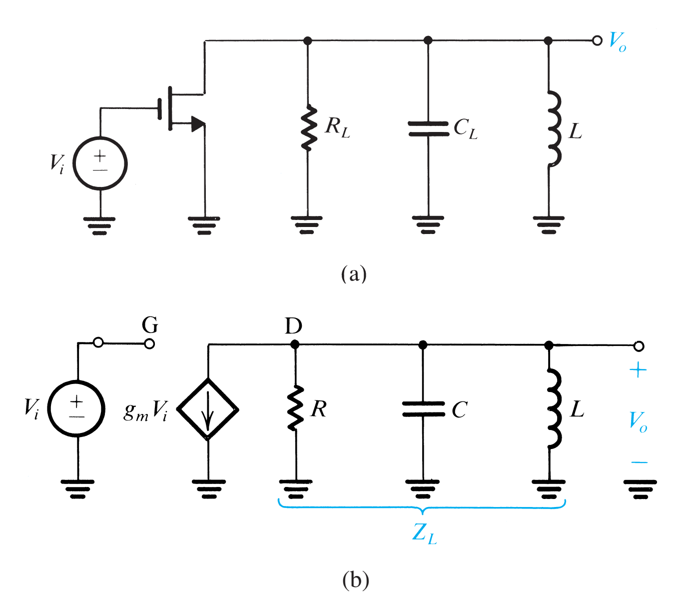


Figure 17.42 The basic principle of tuned amplifiers is illustrated using a MOSFET with a tuned-circuit load. Bias details are not shown.

Thus the voltage gain can be expressed as

$$\frac{V_o}{V_i} = -\frac{g_m}{C}\frac{s}{s^2 + s(1/CR) + 1/LC} \tag{17.114}$$

which is a second-order bandpass function. Thus the tuned amplifier has a center frequency of

$$\omega_0 = 1/\sqrt{LC} \tag{17.115}$$

a 3-dB bandwidth of

$$B = \frac{1}{CR} \tag{17.116}$$

a Q factor of

$$Q \equiv \omega_0/B = \omega_0 CR \tag{17.117}$$

and a center-frequency gain of

$$\frac{V_o(j\omega_0)}{V_i(j\omega_0)} = -g_m R \tag{17.118}$$

Note that the expression for the center-frequency gain could have been written by inspection; at resonance, the reactances of L and C cancel out and the impedance of the parallel LCR circuit reduces to R.

Example 17.5

It is required to design a tuned amplifier of the type shown in Fig. 17.42, having $f_0 = 1$ MHz, 3-dB bandwidth = 10 kHz, and center-frequency gain = –10 V/V. The FET available has at the bias point $g_m = 5$ mA/V and $r_o = 10$ kΩ. The output capacitance is negligibly small. Determine the values of R_L, C_L, and L.

Solution

Center-frequency gain = –10 = –5R. Thus $R = 2$ kΩ. Since $R = R_L \parallel r_o$, then $R_L = 2.5$ kΩ.

$$B = 2\pi \times 10^4 = \frac{1}{CR}$$

Thus

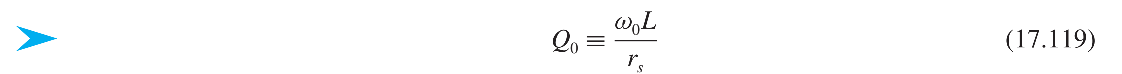
$$C = \frac{1}{2\pi \times 10^4 \times 2 \times 10^3} = 7958 \text{ pF}$$

Since $\omega_0 = 2\pi \times 10^6 = 1/\sqrt{LC}$, we obtain

$$L = \frac{1}{4\pi^2 \times 10^{12} \times 7958 \times 10^{-12}} = 3.18\ \mu\text{H}$$

17.12.2 Inductor Losses

The power loss in the inductor is usually represented by a series resistance r_s as shown in Fig. 17.43(a). However, rather than specifying the value of r_s, the usual practice is to specify the inductor Q factor at the frequency of interest,



$$Q_0 \equiv \frac{\omega_0 L}{r_s} \tag{17.119}$$

Typically, Q_0 is in the range of 50 to 200.

The analysis of a tuned amplifier is greatly simplified by representing the inductor loss by a parallel resistance R_p, as shown in Fig. 17.43(b). The relationship between R_p and Q_0 can be found by writing, for the admittance of the circuit in Fig. 17.43(a),

$$Y(j\omega_0) = \frac{1}{r_s + j\omega_0 L}$$

$$= \frac{1}{j\omega_0 L}\frac{1}{1 - j(1/Q_0)} = \frac{1}{j\omega_0 L}\frac{1 + j(1/Q_0)}{1 + \left(1/Q_0^2\right)}$$

For $Q_0 \gg 1$,

$$Y(j\omega_0) \simeq \frac{1}{j\omega_0 L}\left(1 + j\frac{1}{Q_0}\right) \tag{17.120}$$

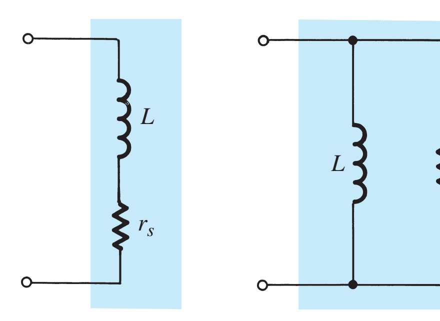


Figure 17.43 Inductor equivalent circuits.

Equating this to the admittance of the circuit in Fig. 17.43(b) gives

$$Q_0 = \frac{R_p}{\omega_0 L} \tag{17.121}$$

or, equivalently,

$$R_p = \omega_0 L Q_0 \tag{17.122}$$

Finally, it should be noted that the coil Q factor poses an upper limit on the value of Q achieved by the tuned circuit.

EXERCISE

17.33 If the inductor in Example 17.5 has $Q_0 = 150$, find R_p and then find the value to which R_L should be changed to keep the overall Q, and hence the bandwidth, unchanged.
Ans. 3 kΩ; 15 kΩ

17.12.3 Use of Transformers

In many cases it is found that the required value of inductance is not practical, in the sense that coils with the required inductance might not be available with the required high values of Q_0. A simple solution is to use a transformer to effect an impedance change. Alternatively, a tapped coil, known as an **autotransformer**, can be used, as shown in Fig. 17.44. Provided the two parts of the inductor are tightly coupled, which can be achieved by winding on a ferrite core, the transformation relationships shown hold. The result is that the tuned circuit seen between terminals 1 and 1′ is equivalent to that in Fig. 17.42(b). For example, if a turns ratio $n = 3$ is used in the amplifier of Example 17.5, then a coil with inductance $L' = 9 \times 3.18 = 28.6\ \mu\text{H}$ and a capacitance $C' = 7958/9 = 884$ pF will be required. Both these values are more practical than the original ones.

In applications that involve coupling the output of a tuned amplifier to the input of another amplifier, the tapped coil can be used to raise the effective input resistance of the latter amplifier stage. In this way, one can avoid reduction of the overall Q. This point is illustrated in Fig. 17.45 and in the following exercises.

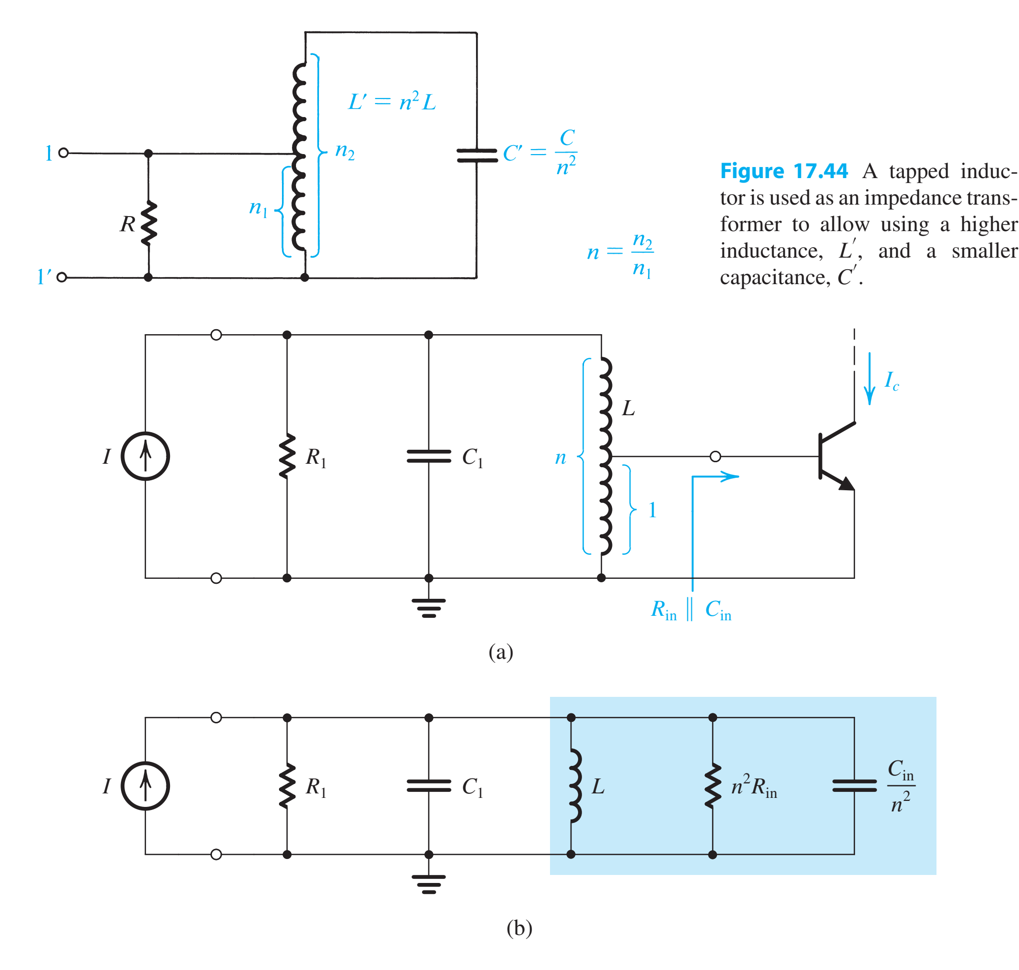


Figure 17.44 A tapped inductor is used as an impedance transformer to allow using a higher inductance, L', and a smaller capacitance, C'.

Figure 17.45 **(a)** The output of a tuned amplifier is coupled to the input of another amplifier via a tapped coil. **(b)** An equivalent circuit. Note that the use of a tapped coil increases the effective input impedance of the second amplifier stage.

EXERCISES

D17.34 Consider the circuit in Fig. 17.45(a), first without tapping the coil. Let $L = 5\ \mu$H and assume that R_1 is fixed at 1 kΩ. We wish to design a tuned amplifier with $f_0 = 455$ kHz and a 3-dB bandwidth of 10 kHz [this is the intermediate frequency (IF) amplifier of an AM radio]. If the BJT has $R_{in} = 1$ kΩ and $C_{in} = 200$ pF, find the actual bandwidth obtained and the required value of C_1.
Ans. 13 kHz; 24.27 nF

D17.35 Since the bandwidth realized in Exercise 17.34 is greater than desired, find an alternative design utilizing a tapped coil as in Fig. 17.45(a). Find the value of n that allows the specifications to be just met. Also find the new required value of C_1 and the current gain I_c/I at resonance. Assume that at the bias point the BJT has $g_m = 40$ mA/V.
Ans. 1.36; 24.36 nF; 19.1 A/A

17.12.4 Amplifiers with Multiple Tuned Circuits

The selectivity achieved with the single tuned circuit of Fig. 17.42 is not sufficient in many applications—for instance, in the IF amplifier of a radio or a TV receiver. Greater selectivity is obtained by using additional tuned stages. Figure 17.46 shows a BJT with tuned circuits at both the input and the output.[12] In this circuit the bias details are shown, from which we note that biasing is quite similar to the classical arrangement employed in low-frequency, discrete-circuit design. However, to avoid the loading effect of the bias resistors R_{B1} and R_{B2} on the input tuned circuit, a **radio-frequency choke** (RFC) is inserted in series with each resistor. Such chokes have low resistance but high impedances at the frequencies of interest. The use of RFCs in biasing tuned RF amplifiers is common practice.

The analysis and design of the double-tuned amplifier of Fig. 17.46 is complicated by the Miller effect[13] due to capacitance C_μ. Since the load is not simply resistive, as was the case in the amplifiers studied in Section 10.3.3, the Miller impedance at the input will be complex. This reflected impedance will cause detuning of the input circuit as well as "skewing" of the response of the input circuit. Needless to say, the coupling introduced by C_μ makes tuning (or aligning) the amplifier quite difficult. Worse still, the capacitor C_μ can cause oscillations to occur (see Gray and Searle, 1969, and Problem 17.101).

Methods exist for **neutralizing** the effect of C_μ, using additional circuits arranged to feed back a current equal and opposite to that through C_μ. An alternative, and preferred, approach is to use circuit configurations that do not suffer from the Miller effect. These are discussed later. Before leaving this section, however, we wish to point out that circuits of the type shown in Fig. 17.46 are usually designed utilizing the *y*-parameter model of the BJT (see

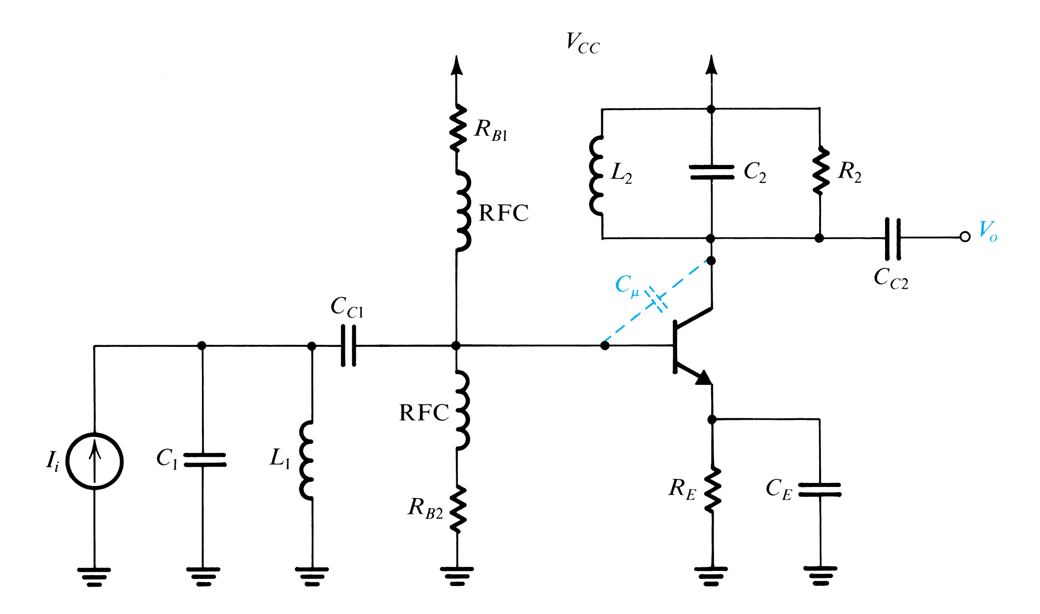


Figure 17.46 A BJT amplifier with tuned circuits at the input and the output.

[12] Note that because the input circuit is a parallel resonant circuit, an input current source (rather than voltage source) signal is utilized.

[13] Here we use "Miller effect" to refer to the effect of the feedback capacitance C_μ in reflecting back an input impedance that is a function of the amplifier load impedance.

Appendix C). This is done because here, in view of the fact that C_μ plays a significant role, the y-parameter model makes the analysis simpler (in comparison to that using the hybrid-π model). Also, the y parameters can easily be measured at the particular frequency of interest, ω_0. For narrow-band amplifiers, the assumption is usually made that the y parameters remain approximately constant over the passband.

17.12.5 The Cascode and the CC–CB Cascade

From our study of amplifier frequency response in Chapter 10, we know that two amplifier configurations do not suffer from the Miller effect. These are the cascode configuration and the common-collector, common-base cascade. Figure 17.47 shows tuned amplifiers based on these two configurations. The CC–CB cascade is usually preferred in IC implementations because its differential structure makes it suitable for IC biasing techniques. (Note that the biasing details of the cascode circuit are not shown in Fig. 17.47(a). Biasing can be done using arrangements similar to those discussed in earlier chapters.)

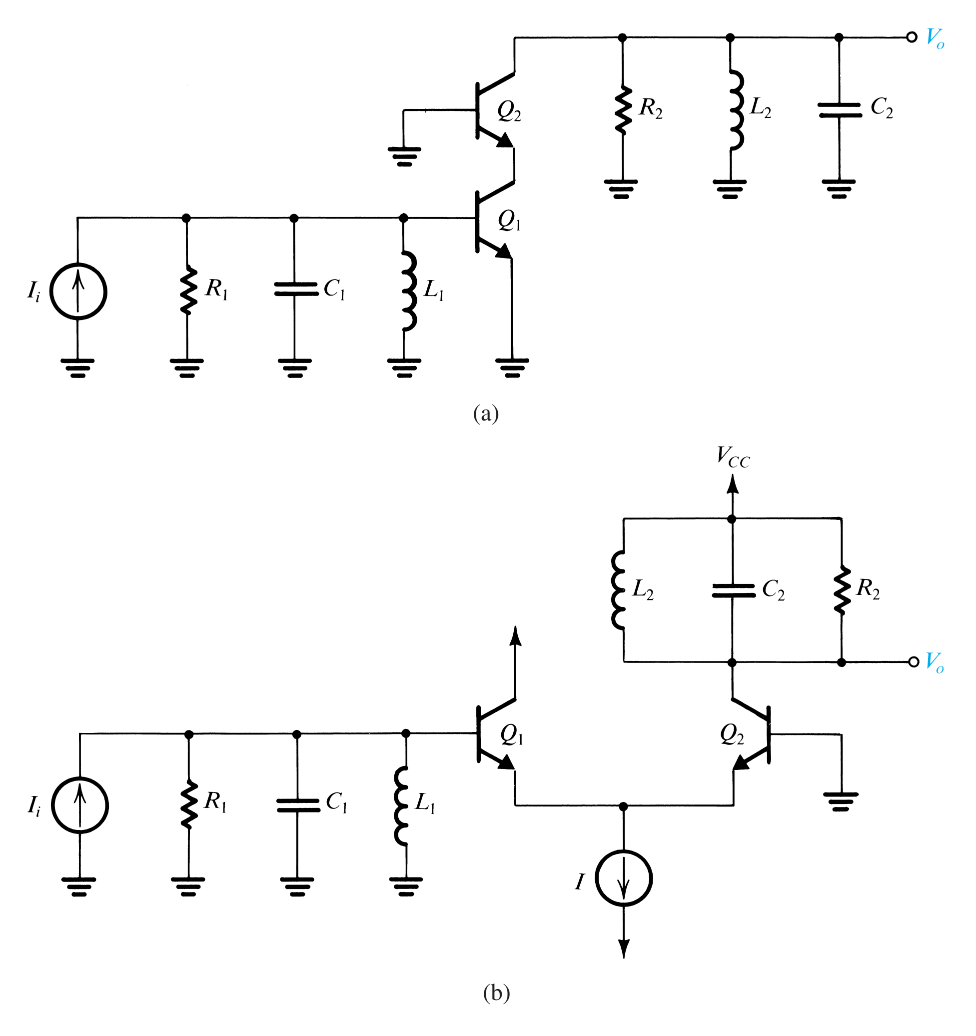


Figure 17.47 Two tuned-amplifier configurations that do not suffer from the Miller effect: **(a)** cascode and **(b)** common-collector, common-base cascade. (Note that bias details of the cascode circuit are not shown.)

17.12.6 Synchronous Tuning and Stagger Tuning

In the design of a tuned amplifier with multiple tuned circuits, the question of the frequency to which each circuit should be tuned arises. The objective, of course, is for the overall response to exhibit high passband flatness and skirt selectivity. To investigate this question, we shall assume that the overall response is the product of the individual responses: in other words, that the stages do not interact. This can easily be achieved using circuits such as those in Fig. 17.47.

Consider first the case of N identical resonant circuits, known as the **synchronously tuned** case. Figure 17.48 shows the response of an individual stage and that of the cascade. Observe the bandwidth "shrinkage" of the overall response. The 3-dB bandwidth B of the overall amplifier is related to that of the individual tuned circuits, ω_0/Q, by (see Problem 17.102)

$$B = \frac{\omega_0}{Q}\sqrt{2^{1/N} - 1} \tag{17.123}$$

The factor $\sqrt{2^{1/N} - 1}$ is known as the **bandwidth-shrinkage factor**. Given B and N, we can use Eq. (17.123) to determine the bandwidth required of the individual stages, ω_0/Q.

EXERCISE

D17.36 Consider the design of an IF amplifier for an FM radio receiver. Using two synchronously tuned stages with $f_0 = 10.7$ MHz, find the 3-dB bandwidth of each stage so that the overall bandwidth is 200 kHz. Using 3-μH inductors find C and R for each stage.
Ans. 310.8 kHz; 73.7 pF; 6.95 kΩ

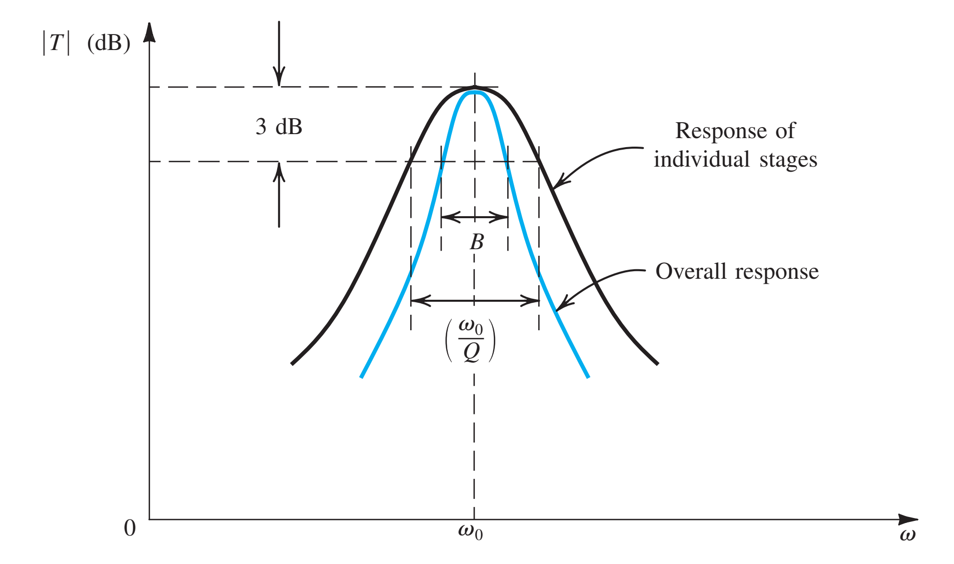


Figure 17.48 Frequency response of a synchronously tuned amplifier.

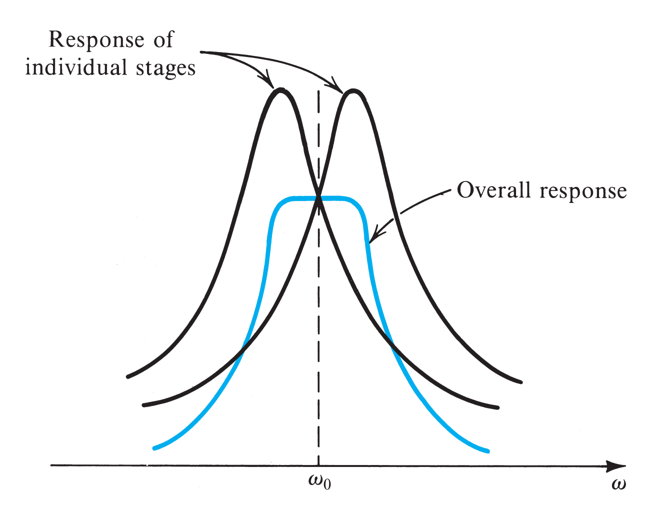


Figure 17.49 Stagger-tuning the individual resonant circuits can result in an overall response with a passband flatter than that obtained with synchronous tuning (Fig. 17.48).

A much better overall response is obtained by stagger-tuning the individual stages, as illustrated in Fig. 17.49. Stagger-tuned amplifiers are usually designed so that the overall response exhibits *maximal flatness* around the center frequency f_0. Such a response can be obtained by transforming the response of a maximally flat (Butterworth) low-pass filter up the frequency axis to ω_0. Appendix H shows how this can be done.

Summary

- A filter is a linear two-port network with a transfer function $T(s) = V_o(s)/V_i(s)$. For physical frequencies, the filter transmission is expressed as $T(j\omega) = |T(j\omega)|e^{j\phi(\omega)}$. The magnitude of transmission can be expressed in decibels using either the gain function $G(\omega) \equiv 20\log|T|$ or the attenuation function $A(\omega) \equiv -20\log|T|$.
- The transmission characteristics of a filter are specified in terms of the edges of the passband(s) and the stopband(s); the maximum allowed variation in passband transmission, A_{max} (dB); and the minimum attenuation required in the stopband, A_{min} (dB). In some applications, the phase characteristics are also specified.
- The filter transfer function can be expressed as the ratio of two polynomials in s; the degree of the denominator polynomial, N, is the filter order. The N roots of the denominator polynomial are the poles (natural modes).
- To obtain a highly selective response, the poles are complex and occur in conjugate pairs (except for one real pole when N is odd). The zeros are placed on the $j\omega$ axis in the stopband(s) including $\omega = 0$ and $\omega = \infty$.
- The Butterworth filter approximation provides a low-pass response that is maximally flat at $\omega = 0$. The transmission decreases monotonically as ω increases, reaching 0 (infinite attenuation) at $\omega = \infty$, where all N transmission zeros lie. Eq. (17.11) gives $|T|$, where ϵ is given by Eq. (17.14) and the order N is determined using Eq. (17.15). The poles are found using the graphical construction of Fig. 17.10, and the transfer function is given by Eq. (17.16).
- The Chebyshev filter approximation provides a low-pass response that is equiripple in the passband with the transmission decreasing monotonically in the stopband. All the transmission zeros are at $s = \infty$. Eq. (17.18) gives $|T|$ in the passband and Eq. (17.19) gives $|T|$ in the stopband, where ϵ is given by Eq. (17.21). The order N can be determined using Eq. (17.22). The poles are given by Eq. (17.23) and the transfer function by Eq. (17.24).
- Figures 17.13 and 17.14 provide a summary of first-order filter functions and their realizations.
- Figure 17.16 provides the characteristics of seven special second-order filtering functions.
- The second-order LCR resonator of Fig. 17.17(a) realizes a pair of complex-conjugate poles with $\omega_0 = 1/\sqrt{LC}$ and $Q = \omega_0 CR$. This resonator can be used to realize the

various special second-order filtering functions, as shown in Fig. 17.18.

- By replacing the inductor of an LCR resonator with a simulated inductance obtained using the Antoniou circuit of Fig. 17.20(a), the op amp–RC resonator of Fig. 17.21(b) is obtained. This resonator can be used to realize the various second-order filter functions as shown in Fig. 17.22. The design equations for these circuits are given in Table 17.1.
- Biquads based on the two-integrator-loop topology are the most versatile and popular second-order filter realizations. There are two varieties: the KHN circuit of Fig. 17.24(a), which realizes the LP, BP, and HP functions simultaneously and can be combined with the output summing amplifier of Fig. 17.24(b) to realize the notch and all-pass functions; and the Tow–Thomas circuit of Fig. 17.25(b), which realizes the BP and LP functions simultaneously. Feedforward can be applied to the Tow–Thomas circuit to obtain the circuit of Fig. 17.26, which can be designed to realize any of the second-order functions (see Table 17.2).
- Single-amplifier biquads (SABs) are obtained by placing a bridged-T network in the negative-feedback path of an op amp. If the op amp is ideal, the poles realized are at the same locations as the zeros of the RC network. The complementary transformation can be applied to the feedback loop to obtain another feedback loop having identical poles. Different transmission zeros are realized by feeding the input signal to circuit nodes that are connected to ground. SABs are economic in their use of op amps but are sensitive to the op-amp nonidealities and are thus limited to low-Q applications ($Q \leq 10$).
- The classical sensitivity function

$$S_x^y = \frac{\partial y/y}{\partial x/x}$$

is a very useful tool in investigating how tolerant a filter circuit is to the unavoidable inaccuracies in component values and to the nonidealities of the op amps.
- Transconductance-C circuits utilize transconductors and capacitors to realize medium- and high-frequency filters (as high as hundreds of megahertz) that can be implemented in CMOS. The basic building block is the integrator, and the basic filter building block is based on the two-integrator-loop topology.
- Switched-capacitor (SC) filters are based on the principle that a capacitor C, periodically switched between two circuit nodes at a high rate, f_c, is equivalent to a resistance $R = 1/Cf_c$ connecting the two circuit nodes. SC filters can be fabricated in monolithic form using CMOS IC technology.
- Tuned amplifiers utilize LC-tuned circuits as loads, or at the input, of transistor amplifiers. They are used in the design of the RF tuner and the IF amplifier of communication receivers. The cascode and the CC–CB cascade configurations are frequently used in the design of tuned amplifiers. Stagger-tuning the individual tuned circuits results in a flatter passband response (in comparison to that obtained with all the resonant circuits synchronously tuned).

PROBLEMS

Section 17.1: Filter Transmission, Types, and Specification

17.1 The transfer function of a first-order low-pass filter (such as that realized by an RC circuit) can be expressed as $T(s) = \omega_0/(s+\omega_0)$, where ω_0 is the 3-dB frequency of the filter. Give in table form the values of $|T|$, ϕ, G, and A at $\omega = 0,\ 0.5\omega_0, \omega_0, 2\omega_0, 5\omega_0, 10\omega_0$, and $100\omega_0$.

17.2 A sinusoid with 1-V peak amplitude is applied at the input of a filter having the transfer function

$$T(s) = \frac{2\pi \times 10^4}{s + 2\pi \times 10^4}$$

Find the peak amplitude and the phase (relative to that of the input sinusoid) of the output sinusoid if the frequency of the input sinusoid is (a) 1 kHz, (b) 10 kHz, (c) 100 kHz, and (d) 1 MHz.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

***17.3** A filter has the transfer function $T(s) = 1/[(s+1)(s^2 + s+1)]$. Show that $|T| = \sqrt{1+\omega^6}$ and find an expression for its phase response $\phi(\omega)$. Calculate the values of $|T|$ and ϕ for $\omega = 0.1$, 1, and 10 rad/s and then find the output corresponding to each of the following input signals:

(a) 10 sin 0.1t (volts)
(b) 10 sin t (volts)
(c) 10 sin 10t (volts)

17.4 For the filter whose magnitude response is sketched (as the blue curve) in Fig. 17.3, find $|T|$ at $\omega = 0$, $\omega = \omega_p$, and $\omega = \omega_s$. $A_{max} = 0.2$ dB, and $A_{min} = 60$ dB.

D 17.5 A low-pass filter is required to pass all signals within its passband, extending from 0 to 4 kHz, with a transmission variation of at most 5% (i.e., the ratio of the maximum to minimum transmission in the passband should not exceed 1.05). The transmission in the stopband, which extends from 5 kHz to ∞, should not exceed 0.05% of the maximum passband transmission. What are the values of A_{max}, A_{min}, and the selectivity factor for this filter?

17.6 A low-pass filter is specified to have $f_p = 5$ kHz and a selectivity factor of 10. The specifications are just met by a first-order transfer function

$$T(s) = \frac{2\pi \times 10^4}{s + 2\pi \times 10^4}$$

What must A_{max} and A_{min} be?

17.7 A low-pass filter is specified to have $A_{max} = 2$ dB and $A_{min} = 12$ dB. It is found that these specifications can be just met with a single-time-constant RC circuit having a time constant of 1 s and a dc transmission of unity. What must ω_p and ω_s of this filter be? What is the selectivity factor?

17.8 Sketch transmission specifications for a high-pass filter having a passband defined by $f \geq 3$ kHz and a stopband defined by $f \leq 2$ kHz. $A_{max} = 0.4$ dB, and $A_{min} = 60$ dB.

17.9 Sketch transmission specifications for a bandstop filter that is required to pass signals over the bands $0 \leq f \leq 10$ kHz and 20 kHz $\leq f \leq \infty$ with A_{max} of 0.5 dB. The stopband extends from $f = 12$ kHz to $f = 18$ kHz, with a minimum required attenuation of 50 dB.

Section 17.2: The Filter Transfer Function

17.10 Consider a fifth-order filter whose poles are all at a radial distance from the origin of 10^4 rad/s. One pair of complex-conjugate poles is at 18° angles from the $j\omega$ axis, and the other pair is at 54° angles. Give the transfer function in each of the following cases.

(a) The transmission zeros are all at $s = \infty$ and the dc gain is unity.
(b) The transmission zeros are all at $s = 0$ and the high-frequency gain is unity.

What type of filter results in each case?

17.11 A third-order low-pass filter has transmission zeros at $\omega = 2$ rad/s and $\omega = \infty$. Its natural modes are at $s = -1$ and $s = -0.5 \pm j0.8$. The dc gain is unity. Find $T(s)$.

17.12 A second-order low-pass filter has poles at $-0.25 \pm j$ and a transmission zero at $\omega = 2$ rad/s. If the dc gain is unity, give the transfer function $T(s)$. What is the gain at ω approaching infinity?

17.13 Find the order N and the form of $T(s)$ of a bandpass filter having transmission zeros as follows: one at $\omega = 0$, one at $\omega = 10^3$ rad/s, one at 3×10^3 rad/s, one at 6×10^3 rad/s, and one at $\omega = \infty$. If this filter has a monotonically decreasing passband transmission with a peak at the center frequency of 2×10^3 rad/s, and equiripple response in the stopbands, sketch the shape of its $|T|$.

***17.14** Analyze the RLC network of Fig. P17.14 to determine its transfer function $V_o(s)/V_i(s)$ and hence its poles and zeros. (*Hint:* Begin the analysis at the output and work your way back to the input.)

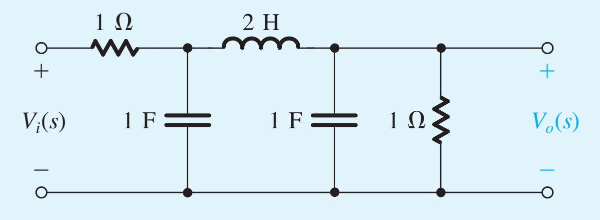


Figure P17.14

Section 17.3: Butterworth and Chebyshev Filters

D 17.15 Determine the order N of the Butterworth filter for which $A_{max} = 0.5$ dB, $A_{min} \geq 20$ dB, and the selectivity ratio $\omega_s/\omega_p = 1.7$. What is the actual value of minimum stopband attenuation realized? If A_{min} is to be exactly 20 dB, to what value can A_{max} be reduced?

17.16 Show that the order N of a Butterworth filter can be obtained from the approximate expression

$$N \geq \frac{A_{min} - 20 \log \epsilon}{20 \log(\omega_s/\omega_p)}$$

Hint: Use Eq. (17.15) and neglect the unity term.

17.17 Calculate the value of attenuation obtained at a frequency 1.8 times the 3-dB frequency of a seventh-order Butterworth filter.

17.18 Find the natural modes of a Butterworth filter having a 0.5-dB bandwidth of 10^3 rad/s and $N = 5$.

D 17.19 Design a Butterworth filter that meets the following low-pass specifications: $f_p = 10$ kHz, $A_{max} = 3$ dB, $f_s = 20$ kHz, and $A_{min} = 20$ dB. Find N, the natural modes, and $T(s)$. What is the attenuation provided at 30 kHz?

17.20 Sketch the transfer function magnitude for a low-pass Chebyshev filter of (a) sixth order and (b) seventh order.

17.21 On the same diagram, sketch the magnitude of the transfer function of a Butterworth and a Chebyshev low-pass filter of fifth order and having the same ω_p and A_{max}. At the stopband edge, ω_s, which filter gives greater attenuation?

***17.22** Sketch $|T|$ for a seventh-order low-pass Chebyshev filter with $\omega_p = 1$ rad/s and $A_{max} = 0.5$ dB. Use Eq. (17.18) to determine the values of ω at which $|T| = 1$ and the values of ω at which $|T| = 1/\sqrt{1+\epsilon^2}$. Indicate these values on your sketch. Use Eq. (17.19) to determine $|T|$ at $\omega = 2$ rad/s, and indicate this point on your sketch. For large values of ω, at what rate (in dB/octave) does the transmission decrease?

17.23 Contrast the attenuation provided by a sixth-order Chebyshev filter at $\omega_s = 2\omega_p$ to that provided by a Butterworth filter of equal order. For both, $A_{max} = 1$ dB. Sketch $|T|$ for both filters on the same axes.

D *17.24 It is required to design a low-pass filter to meet the following specifications: $f_p = 3.4$ kHz, $A_{max} = 1$ dB, $f_s = 4$ kHz, $A_{min} = 35$ dB.

(a) Find the required order of Chebyshev filter. What is the excess (above 35 dB) stopband attenuation obtained?
(b) Find the poles and the transfer function.

Section 17.4: First-Order and Second-Order Filter Functions

D 17.25 Use the information displayed in Fig. 17.13 to design a first-order op amp–RC low-pass filter having a 3-dB frequency of 5 kHz, a dc gain magnitude of 10, and an input resistance of 12 kΩ.

D 17.26 Use the information given in Fig. 17.13 to design a first-order op amp–RC high-pass filter with a 3-dB frequency of 200 Hz, a high-frequency input resistance of 120 kΩ, and a high-frequency gain magnitude of unity.

17.27 Derive an expression for the transfer function of the op amp–RC circuit that is shown in Fig. 17.13(c). Give expressions for the frequency of the transmission zero ω_Z, the frequency of the pole ω_P, the dc gain, and the high-frequency gain.

D *17.28 Use the information given in Fig. 17.13 to design a first-order op amp–RC spectrum-shaping network with a transmission zero frequency of 100 Hz, a pole frequency of 10 kHz, and a dc gain magnitude of unity. The low-frequency input resistance is to be 10 kΩ. What is the high-frequency gain that results? Sketch the magnitude of the transfer function versus frequency.

D *17.29 By cascading a first-order op amp–RC low-pass circuit with a first-order op amp–RC high-pass circuit, one can design a wideband bandpass filter. Provide such a design for the case in which the midband gain is 12 dB and the 3-dB bandwidth extends from 50 Hz to 50 kHz. Select appropriate component values under the constraints that no resistors higher than 100 kΩ are to be used and that the input resistance is to be as high as possible.

D 17.30 Derive $T(s)$ for the op amp–RC circuit in Fig. 17.14. Find $|T(j\omega)|$ and $\phi(\omega)$ We wish to use this circuit as a variable phase shifter by adjusting R. If the input signal frequency is 5×10^3 rad/s and if $C = 10$ nF, find the values of R required to obtain phase shifts of –30°, –60°, –90°, –120°, and –150°.

17.31 Show that by interchanging R and C in the op amp–RC circuit of Fig. 17.14, the resulting phase shift covers the range 0 to 180° (with 0° at high frequencies and 180° at low frequencies).

D *17.32 Use two first-order op amp–RC all-pass circuits in cascade to design a circuit that provides a set of three-phase 60-Hz voltages, each separated by 120° and equal in magnitude, as shown in the phasor diagram of Fig. P17.32. These voltages simulate those used in three-phase power transmission systems. Use 1-µF capacitors.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

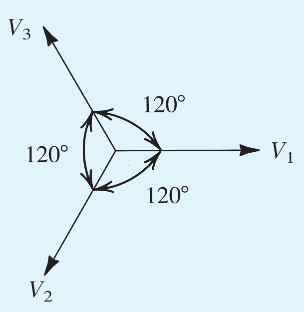


Figure P17.32

17.33 Use the information in Fig. 17.16(a) to obtain the transfer function of a second-order low-pass filter with $\omega_0 = 10^4$ rad/s, $Q = 2$, and dc gain $= 1$. At what frequency does $|T|$ peak? What is the peak transmission?

D *17.34 Use the information in Fig. 17.16(a) to obtain the transfer function of a second-order low-pass filter that just meets the specifications defined in Fig. 17.3 with $\omega_p = 1$ rad/s and $A_{max} = 3$ dB. Note that there are two possible solutions. For each, find ω_0 and Q. Also, if $\omega_s = 2$ rad/s, find the value of A_{min} obtained in each case.

17.35 Find the transfer function of a second-order high-pass filter with a maximally flat passband response, a 3-dB frequency at $\omega = 1$ rad/s, and a high-frequency gain of unity. Give the location of the poles and zeros.

17.36 Use the information given in Fig. 17.16(b) to find the transfer function of a second-order high-pass filter with natural modes at $-0.5 \pm j\sqrt{3}/2$ and a high-frequency gain of unity. What are ω_0 and Q of the poles?

17.37 Find the transfer function of a second-order bandpass filter for which the center frequency $f_0 = 10$ kHz, the 3-dB bandwidth is 500 Hz, and the center-frequency gain is 10. Also, give the locations of the poles and zeros.

D **17.38 (a) Show that $|T|$ of a second-order bandpass function is geometrically symmetrical around the center frequency ω_0. That is, the members of each pair of frequencies ω_1 and ω_2 for which $|T(j\omega_1)| = |T(j\omega_2)|$ are related by $\omega_1\omega_2 = \omega_0^2$.

(b) Find the transfer function of the second-order bandpass filter that meets specifications of the form in Fig. 17.4 where $\omega_{p1} = 8100$ rad/s, $\omega_{p2} = 10{,}000$ rad/s, and $A_{max} = 3$ dB. If $\omega_{s1} = 3000$ rad/s find A_{min} and ω_{s2}.

D *17.39 Use the result of Exercise 17.15 to find the transfer function of a notch filter that is required to eliminate a bothersome interference of 60-Hz frequency. Since the frequency of the interference is not stable, the filter should be designed to provide attenuation ≥ 20 dB over a 6-Hz band centered around 60 Hz. The dc transmission of the filter is to be unity.

17.40 Consider a second-order all-pass circuit in which errors in the component values result in the frequency of the zeros being slightly lower than that of the poles. Roughly sketch the expected $|T|$. Repeat for the case of the frequency of the zeros slightly higher than the frequency of the poles.

17.41 Consider a second-order all-pass filter in which errors in the component values result in the Q factor of the zeros being greater than the Q factor of the poles. Roughly sketch the expected $|T|$. Repeat for the case of the Q factor of the zeros lower than the Q factor of the poles.

Section 17.5: The Second-Order LCR Resonator

17.42 Analyze the circuit in Fig. 17.17(c) to determine its transfer function $T(s) \equiv V_o(s)/V_i(s)$, and hence show that its poles are characterized by ω_0 and Q of Eqs. (17.34) and (17.35), respectively.

D 17.43 Design the LCR resonator of Fig. 17.17(a) to obtain natural modes with $\omega_0 = 10^5$ rad/s and $Q = 5$. Use $R = 10$ kΩ.

17.44 For the LCR resonator of Fig. 17.17(a), find the change in ω_0 that results from

(a) increasing L by 1%
(b) increasing C by 1%
(c) decreasing R by 1%

17.45 For each of the circuits in Fig. P17.45, find the transmission as ω approaches zero and as ω approaches ∞, and hence find the transmission zeros.

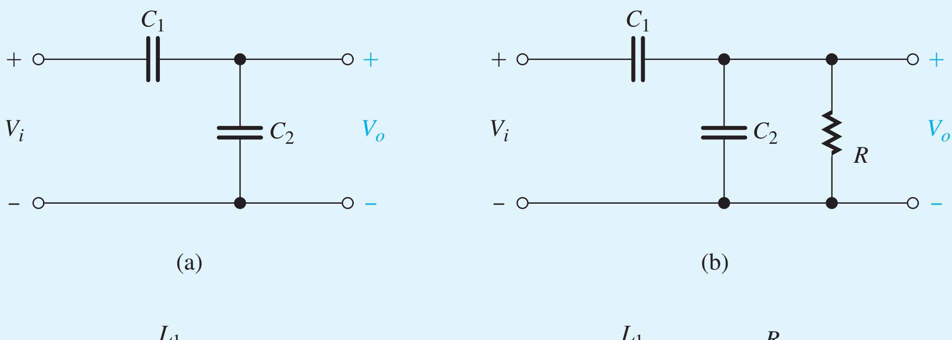
17.46 Derive an expression for $V_o(s)/V_i(s)$ of the high-pass circuit in Fig. 17.18(c).

D 17.47 Use the circuit of Fig. 17.18(b) to design a low-pass filter with $\omega_0 = 10^6$ rad/s and $Q = 1/\sqrt{2}$. Utilize a 1-nF capacitor.

D 17.48 Modify the bandpass circuit of Fig. 17.18(d) to change its center-frequency gain from 1 to 0.5 without changing ω_0 or Q.

17.49 Consider the LCR resonator of Fig. 17.17(a) with node x disconnected from ground and connected to an input signal

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem



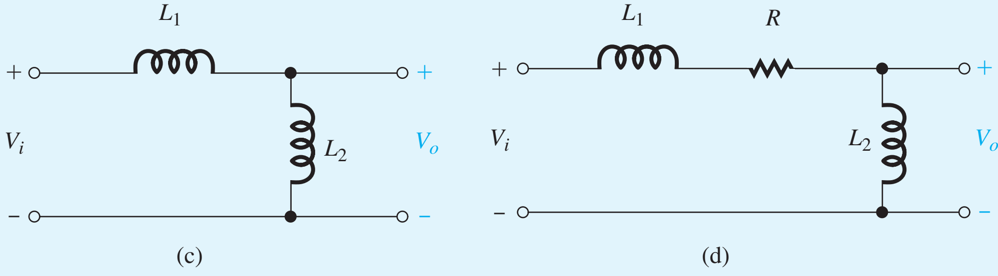


Figure P17.45

source V_x, node y disconnected from ground and connected to another input signal source V_y, and node z disconnected from ground and connected to a third input signal source V_z. Use superposition to find the voltage that develops across the resonator, V_o, in terms of V_x, V_y, and V_z.

17.50 Consider the notch circuit shown in Fig. 17.18(g). For what ratio of C_2 to C_1 does the notch occur at 1.1 ω_0? For this case, what is the magnitude of the transmission at frequencies $\ll \omega_0$? At frequencies $\gg \omega_0$?

Section 17.6: Second-Order Active Filters Based on Inductor Replacement

D 17.51 Design the circuit of Fig. 17.20 (utilizing suitable component values) to realize an inductance of (a) 15 H, (b) 1.5 H, and (c) 0.15 H.

17.52 Figure P17.52 shows a generalized form of the Antoniou circuit of Fig. 17.20(a). Here, R_5 is eliminated and the other four components are replaced by general impedances Z_1, Z_2, Z_3, and Z_4.

(a) With an impedance Z_5 connected between node 2 and ground, show that the input impedance looking into port 1 (i.e., between node 1 and ground) is

$$Z_{11} = \left(\frac{Z_1 Z_3}{Z_2 Z_4}\right) Z_5$$

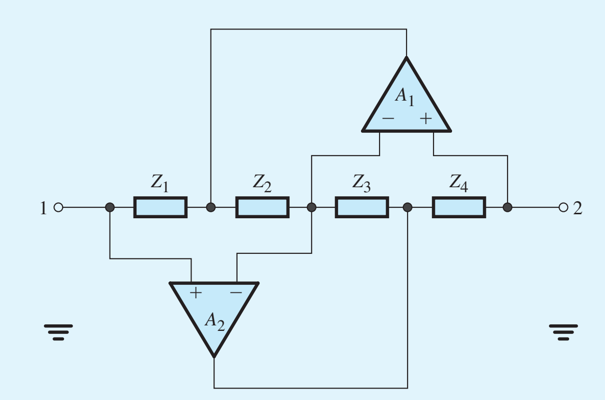


Figure P17.52

(b) From the symmetry of the circuit, show that if an impedance Z_6 is connected between terminal 1 and ground, the input impedance looking into port 2, which is between terminal 2 and ground, is given by

$$Z_{22} = \left(\frac{Z_2 Z_4}{Z_1 Z_3}\right) Z_6$$

(c) From the expressions above, observe that the two-port network in Fig. P17.52 acts as an "impedance transformer." Since by the appropriate choice of Z_1, Z_2, Z_3, and Z_4, the transformation ratio can be a general function

of the complex frequency variable s, the circuit is known as a generalized impedance converter, or GIC.

17.53 Consider the Antoniou circuit of Fig. 17.20(a) with R_5 eliminated, a capacitor C_6 connected between node 1 and ground, and a voltage source V_2 connected to node 2. Show that the input impedance seen by V_2 is $R_2/s^2C_4C_6R_1R_3$. How does this impedance behave for physical frequencies ($s = j\omega$)? (This impedance is known as a **frequency-dependent negative resistance**, or FDNR.)

***17.54** Starting from first principles and assuming ideal op amps, derive the transfer function of the circuit in Fig. 17.22(a).

D *17.55 It is required to design a fifth-order Butterworth filter having a 3-dB bandwidth of 10^5 rad/s and a unity dc gain. Use a cascade of two circuits of the type shown in Fig. 17.22(a) and a first-order op amp–RC circuit of the type shown in Fig. 17.13(a). Select appropriate component values.

D 17.56 Design the circuit of Fig. 17.22(e) to realize an LPN function with $f_0 = 10$ kHz, $f_n = 12$ kHz, $Q = 10$, and a unity dc gain. Select $C_4 = 10$ nF.

D 17.57 Design the all-pass circuit of Fig. 17.22(g) to provide a phase shift of 180° at $f = 2$ kHz and to have $Q = 2$. Use 1-nF capacitors.

D 17.58 Using the transfer function of the LPN filter, given in Table 17.1, derive the design equations also given.

D 17.59 Using the transfer function of the HPN filter, given in Table 17.1, derive the design equations also given.

D **17.60 It is required to design a third-order low-pass filter whose $|T|$ is equiripple in both the passband and the stopband (in the manner shown in Fig. 17.3, except that the response shown is for $N = 5$). The filter passband extends from $\omega = 0$ to $\omega = 1$ rad/s, and the passband transmission varies between 1 and 0.9. The stopband edge is at $\omega = 1.2$ rad/s. The following transfer function was obtained using filter-design tables:

$$T(s) = \frac{0.4508\left(s^2 + 1.6996\right)}{(s + 0.7294)(s^2 + s0.2786 + 1.0504)}$$

The actual filter realized is to have $\omega_p = 10^5$ rad/s.

(a) Obtain the transfer function of the actual filter by replacing s by $s/10^5$.

(b) Realize this filter as the cascade connection of a first-order LP op amp–RC circuit of the type shown in Fig. 17.13(a) and a second-order LPN circuit of the type shown in Fig. 17.22(e). Each section is to have a dc gain of unity. Select appropriate component values. (*Note:* A filter with an equiripple response in both the passband and the stopband is known as an **elliptic filter**.)

Section 17.7: Second-Order Active Filters Based on the Two-Integrator-Loop Topology

D 17.61 Design the KHN circuit of Fig. 17.24(a) to realize a bandpass filter with a center frequency of 2 kHz and a 3-dB bandwidth of 50 Hz. Use 10-nF capacitors. Give the complete circuit and specify all component values. What value of center-frequency gain is obtained?

D 17.62 (a) Using the KHN biquad with the output summing amplifier of Fig. 17.24(b), show that an all-pass function is realized by selecting $R_L = R_H = R_B/Q$. Also show that the flat gain obtained is KR_F/R_H.
(b) Design the all-pass circuit to obtain $\omega_0 = 10^5$ rad/s, $Q = 4$, and flat gain = 10. Select appropriate component values.

17.63 Consider the case of the KHN circuit used together with the summing amplifier in Fig. 17.24(b) to realize a notch filter with a notch frequency ω_n and a high-frequency gain of G. Find expressions for the values required of the resistances associated with the summing amplifier.

D 17.64 Consider a notch filter with $\omega_n = \omega_0$ realized by using the KHN biquad with an output summing amplifier. If the summing resistors used have 1% tolerances, what is the worst-case percentage deviation between ω_n and ω_0?

D 17.65 Design the circuit of Fig. 17.26 to realize a low-pass notch filter with $\omega_0 = 10^5$ rad/s, $Q = 10$, dc gain = 1, and $\omega_n = 1.3 \times 10^5$ rad/s. Use $C = 10$ nF and $r = 20$ kΩ.

D 17.66 In the all-pass realization using the circuit of Fig. 17.26, which component(s) does one need to trim to adjust (a) only ω_z and (b) only Q_z?

D **17.67 Repeat Problem 17.60 using the Tow–Thomas biquad of Fig. 17.26 to realize the second-order section in the cascade.

Section 17.8: Single-Amplifier Biquadratic Active Filters

D 17.68 Design the circuit of Fig. 17.29 to realize a pair of poles with $\omega_0 = 10^5$ rad/s and $Q = 1/\sqrt{2}$. Use $C_1 = C_2 = 1$ nF.

***17.69** Derive the transfer function $t(s)$ of the bridged-T network in Fig. 17.28(a), and thus verify the expression given in the figure.

17.70 Consider the bridged-T network of Fig. 17.28(a) with $R_3 = R_4 = R$ and $C_1 = C_2 = C$, and denote $CR = \tau$. Find the zeros and poles of the bridged-T network. If the network is placed in the negative-feedback path of an ideal infinite-gain op amp, as in Fig. 17.29, find the poles of the closed-loop amplifier.

17.71 Consider the bridged-T network of Fig. 17.28(b) with $R_1 = R_2 = R$, $C_4 = C$, and $C_3 = C/36$. Let the network be placed in the negative-feedback path of an infinite-gain op amp and let C_4 be disconnected from ground and connected to the input signal source V_i. Analyze the resulting circuit to determine its transfer function $V_o(s)/V_i(s)$, where $V_o(s)$ is the voltage at the op-amp output. Show that the circuit obtained is a bandpass filter and find its ω_0, Q, and the center-frequency gain.

D 17.72 Use the circuit in Fig. 17.30(b) with $\alpha = 1$ to realize a bandpass filter with a center frequency of 10 kHz and a 3-dB bandwidth of 2 kHz. Give the values of all components and specify the center-frequency gain obtained.

D **17.73 Consider the bandpass circuit shown in Fig. 17.30(a). Let $C_1 = C_2 = C$, $R_3 = R$, $R_4 = R/4Q^2$, $CR = 2Q/\omega_0$, and $\alpha = 1$. Disconnect the positive input terminal of the op amp from ground and apply V_i through a voltage divider R_1, R_2 to the positive input terminal as well as through R_4/α as before. Analyze the circuit to find its transfer function V_o/V_i. Find the ratio R_2/R_1 so that the circuit realizes (a) an all-pass function and (b) a notch function. Assume the op amp to be ideal.

D *17.74 Derive the transfer function of the circuit in Fig. 17.33(b) assuming the op amp to be ideal. Thus show that the circuit realizes a high-pass function. What is the high-frequency gain of the circuit? Design the circuit for a maximally flat response with a 3-dB frequency of 10^4 rad/s. Use $C_1 = C_2 = 10$ nF. (*Hint:* For a maximally flat response, $Q = 1/\sqrt{2}$ and $\omega_{3dB} = \omega_0$.)

D *17.75 Design a fifth-order Butterworth low-pass filter that has a 3-dB bandwidth of 10 kHz and a dc gain of unity. Use the cascade connection of two Sallen-and-Key circuits [Fig. 17.34(c)] and a first-order section [Fig. 17.13(a)]. Use a 10-kΩ value for all resistors.

D 17.76 The process of obtaining the complement of a transfer function by interchanging input and ground, as illustrated in Fig. 17.31, applies to any general network (not just RC networks as shown). Show that if the network n is a bandpass with a center-frequency gain of unity, then the complement obtained is a notch. Verify this by using the RLC circuits of Fig. 17.18(d) and (e).

Section 17.9: Sensitivity

17.77 Evaluate the sensitivities of ω_0 and Q relative to R, L, and C of the low-pass circuit in Fig. 17.18(b).

***17.78** Verify the following sensitivity identities:

(a) If $y = uv$, then $S_x^y = S_x^u + S_x^v$.
(b) If $y = u/v$, then $S_x^y = S_x^u - S_x^v$.
(c) If $y = ku$, where k is a constant, then $S_x^y = S_x^u$.
(d) If $y = u^n$, where n is a constant, then $S_x^y = nS_x^u$.
(e) If $y = f_1(u)$ and $u = f_2(x)$, then $S_x^y = S_u^y \cdot S_x^u$.

***17.79** For the high-pass filter of Fig. 17.33(b), what are the sensitivities of ω_0 and Q to amplifier gain A?

***17.80** For the feedback loop of Fig. 17.34(a), use the expressions in Eqs. (17.77) and (17.78) to determine the sensitivities of ω_0 and Q relative to all passive components for the design in which $R_1 = R_2$.

17.81 For the op amp–RC resonator of Fig. 17.21(b), use the expressions for ω_0 and Q given in the top row of Table 17.1 to determine the sensitivities of ω_0 and Q to all resistors and capacitors.

Section 17.10: Transconductance-C Filters

17.82 For the fully differential transconductor of Fig. 17.35(f), find an expression for G_m in terms of I and the MOSFET's transconductance parameter k_n. For $k_n = 0.5$ mA/V, find the bias current I that results in $G_m = 0.25$ mA/V. If for tuning purposes it is required to adjust G_m in a ±5% range, what is the required range for adjusting I?

D 17.83 Using the circuit of Fig. 17.36(a) to realize a 1-kΩ resistance, what G_m is needed? If the output resistance of the transconductor is 100 kΩ, what is the resistance actually realized?

D 17.84 Using four transconductors, give the circuit for obtaining an output voltage V_o related to three input voltages

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

V_1, V_2, and V_3 by $V_o = V_1 - 2V_2 + 3V_3$. Give the values of the four transconductances as ratios of G_m of the transconductor that delivers the output voltage.

D 17.85 For the integrator in Fig. 17.36(b), what value of G_m is needed to obtain an integrator with a unity-gain frequency of 10 MHz utilizing a 5-pF capacitor?

D 17.86 If the transconductor in the integrator of Fig. 17.36(b) has an output resistance R_o and an output capacitance C_o, what is the transfer function realized? If the error in the integrator time constant must be less than 1%, what is the smallest value of C that can be used? If the low-frequency pole introduced by R_o is to be at least two decades lower than the unity-gain frequency of the integrator, what is the smallest G_m that this transconductor must have?

D 17.87 Design the first-order low-pass filter in Fig. 17.36(c) to have a pole frequency of 20 MHz and a dc gain of 10. Use $C = 2$ pF.

D 17.88 If a capacitor C_1 is connected between the input node and node X in the circuit of Fig. 17.36(c), what transfer function V_o/V_i is realized?

D *17.89 For the circuit in Fig. P17.89:

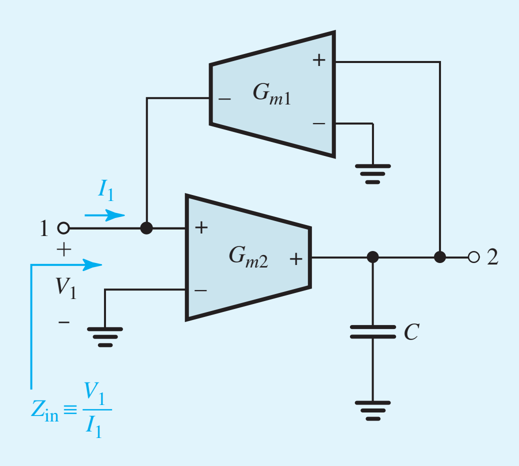


Figure P17.89

(a) Show that the input impedance Z_{in} is that of an inductance L and find an expression for L.

(b) Use the inductance generated at the input to form an LCR resonator. To realize the resistance R, use a third transconductor G_{m3}.

(c) Use a fourth inverting transconductor G_{m4} with its input connected to an input voltage V_i and its output connected to node 1. Compare the circuit thus created to the two-integrator-loop G_m–C filter of Fig. 17.37(b).

(d) What filter function is (i) V_1/V_i, (ii) V_2/V_i?

D 17.90 Consider the G_m–C second-order bandpass filter of Fig. 17.37(b) and its associated expressions in Eqs. (17.95) and (17.96). Generate an alternative design based on selecting $G_{m1} = G_{m2} = G_{m3} = G_m$ and $C_2 = C$. Find expressions for C_1 and G_m in terms of ω_0, Q, and C.

D 17.91 To enable the second-order G_m–C circuit in Fig. 17.37(b) to realize filter functions other than bandpass and lowpass, implement the following modifications:

(a) Connect a capacitor C_3 between the positive input terminal and the output terminal of transconductor G_{m4} and

(b) Add a fifth negative transconductor G_{m5} with V_i applied to its input, and its output connected to the node at which V_2 is taken.

Derive an expression for the transfer function V_1/V_i.

D 17.92 Design the circuit of Fig. 17.37(c) to realize a bandpass function having a center frequency of 25 MHz, $Q = 5$, and a center-frequency gain of 5. Select $G_{m1} = G_{m2}$ and $C_1 = C_2 = 5$ pF.

Section 17.11: Switched-Capacitor Filters

17.93 For the switched-capacitor input circuit of Fig. 17.38(b), in which a clock frequency of 200 kHz is used, what input resistances correspond to capacitance C_1 values of 1 pF, 5 pF, and 10 pF?

17.94 For a dc voltage of 1 V applied to the input of the circuit of Fig. 17.38(b), in which C_1 is 1 pF, what charge is transferred for each cycle of the two-phase clock? For a 100-kHz clock, what is the average current drawn from the input source? For a feedback capacitance of 10 pF, what change would you expect in the output for each cycle of the clock? For an amplifier that saturates at ±10 V and the feedback capacitor initially discharged, how many clock cycles would it take to saturate the amplifier? What is the average slope of the staircase output voltage produced?

D 17.95 Repeat Exercise 17.32 for a clock frequency of 500 kHz.

D 17.96 Repeat Exercise 17.32 for $Q = 50$.

D 17.97 Design the circuit of Fig. 17.40(b) to realize, at the output of the second (noninverting) integrator, a maximally flat low-pass function with $\omega_{3dB} = 10^3$ rad/s and unity dc gain. Use a clock frequency $f_c = 100$ kHz and select $C_1 = C_2 = 5$ pF. Give the values of C_3, C_4, C_5, and C_6. (*Hint:* For a maximally flat response, $Q = 1/\sqrt{2}$ and $\omega_{3dB} = \omega_0$.)

Section 17.12: Tuned Amplifiers

***17.98** A voltage signal source with a resistance $R_s = 10$ kΩ is connected to the input of a common-emitter BJT amplifier. Between base and emitter is connected a tuned circuit with $L = 0.5$ μH and $C = 200$ pF. The transistor is biased at 1 mA and has $\beta = 200$, $C_\pi = 10$ pF, and $C_\mu = 0.5$ pF. The transistor load is a resistance of 5 kΩ. Find ω_0, Q, the 3-dB bandwidth, and the center-frequency gain of this single-tuned amplifier.

17.99 A coil having an inductance of 10 μH is intended for applications around 1-MHz frequency. Its Q is specified to be 250. Find the equivalent parallel resistance R_p. What is the value of the capacitor required to produce resonance at 1 MHz? What additional parallel resistance is required to produce a 3-dB bandwidth of 12 kHz?

17.100 An inductance of 36 μH is resonated with a 1000-pF capacitor. If the inductor is tapped at one-third of its turns and a 1-kΩ resistor is connected across the one-third part, find f_0 and Q of the resonator.

***17.101** Consider a common-emitter transistor amplifier loaded with an inductance L. Ignoring r_o, show that for $\omega C_\mu \ll 1/\omega L$, the amplifier input admittance is given by

$$Y_{in} \simeq \left(\frac{1}{r_\pi} - \omega^2 C_\mu L g_m\right) + j\omega(C_\pi + C_\mu)$$

(*Note:* The real part of the input admittance can be negative. This can lead to oscillations.)

***17.102** (a) Substituting $s = j\omega$ in the transfer function $T(s)$ of a second-order bandpass filter [see Fig. 17.16(c)], find $|T(j\omega)|$. For ω in the vicinity of ω_0 [i.e., $\omega = \omega_0 + \delta\omega = \omega_0(1 + \delta\omega/\omega_0)$, where $\delta\omega/\omega_0 \ll 1$ so that $\omega^2 \simeq \omega_0^2(1 + 2\delta\omega/\omega_0)$], show that, for $Q \gg 1$,

$$|T(j\omega)| \simeq \frac{|T(j\omega_0)|}{\sqrt{1 + 4Q^2(\delta\omega/\omega_0)^2}}$$

(b) Use the result obtained in (a) to show that the 3-dB bandwidth B, of N synchronously tuned sections connected in cascade, is

$$B = (\omega_0/Q)\sqrt{2^{1/N} - 1}$$

****17.103** (a) Using the fact that for $Q \gg 1$ the second-order bandpass response in the neighborhood of ω_0 is the same as the response of a first-order low-pass with 3-dB frequency of $(\omega_0/2Q)$, show that the bandpass response at $\omega = \omega_0 + \delta\omega$, for $\delta\omega \ll \omega_0$, is given by

$$|T(j\omega)| \simeq \frac{|T(j\omega_0)|}{\sqrt{1 + 4Q^2(\delta\omega/\omega_0)^2}}$$

(b) Use the relationship derived in (a) together with Eq. (17.123) to show that a bandpass amplifier with a 3-dB bandwidth B, designed using N synchronously tuned stages, has an overall transfer function given by

$$|T(j\omega)|_{\text{overall}} = \frac{|T(j\omega_0)|_{\text{overall}}}{[1 + 4(2^{1/N} - 1)(\delta\omega/B)^2]^{N/2}}$$

(c) Use the relationship derived in (b) to find the attenuation (in decibels) obtained at a bandwidth $2B$ for $N = 1$ to 5. Also find the ratio of the 30-dB bandwidth to the 3-dB bandwidth for $N = 1$ to 5.

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

CHAPTER 18

Signal Generators and Waveform-Shaping Circuits

IN THIS CHAPTER YOU WILL LEARN

1. That an oscillator circuit that generates sine waves can be implemented by connecting a frequency-selective network in the positive-feedback path of an amplifier.
2. The conditions under which sustained oscillations are obtained and the frequency of the oscillations.
3. How to design nonlinear circuits to control the amplitude of the sine wave obtained in a linear oscillator.
4. A variety of circuits for implementing a linear sine-wave oscillator.
5. How op amps can be combined with resistors and capacitors to implement precision multivibrator circuits.
6. How a bistable circuit can be connected in a feedback loop with an op-amp integrator to implement a generator of square and triangular waveforms.
7. The application of one of the most popular IC chips of all time, the 555 timer, in the design of generators of pulse and square waveforms.
8. How a triangular waveform can be shaped by a nonlinear circuit to provide a sine waveform.

Introduction

In the design of electronic systems, the need frequently arises for signals having prescribed standard waveforms, for example, sinusoidal, square, triangular, or pulse. Systems in which standard signals are required include computer and control systems where clock pulses are needed for, among other things, timing; communication systems where signals of a variety of waveforms are utilized as information carriers; and test and measurement systems where signals, again of a variety of waveforms, are employed for testing and characterizing electronic devices and circuits. In this chapter we study signal-generator circuits.

The signal-generator or oscillator circuits studied in this chapter are collectively capable of providing signals with frequencies in the range of hertz to hundreds of gigahertz. While some can be fabricated on chip, others utilize discrete components. Examples of commonly encountered oscillators include the microprocessor clock generator (fabricated on chip utilizing the ring oscillator studied in Section 16.4.4 with frequencies in the several-gigahertz range); the carrier-waveform generator in wireless transceivers (on chip, up to the hundreds-of-gigahertz range); the oscillator in an electronic watch (utilizing a quartz

crystal with a frequency of 2^{15} Hz); and the variable-frequency function generator in the electronics lab (utilizing a discrete circuit with frequency in the hertz to megahertz range).

There are two distinctly different approaches for the generation of sinusoids, perhaps the most commonly used of the standard waveforms. The first approach, studied in Sections 18.1 to 18.3, employs a **positive-feedback loop** consisting of an amplifier and an RC or LC **frequency-selective network**. While the frequency of the generated sine wave is determined by the frequency-selective network, the amplitude is set using a nonlinear mechanism, implemented either with a separate circuit or using the nonlinearities of the amplifying device itself. In spite of this, these circuits, which generate sine waves utilizing resonance phenomena, are known as **linear oscillators**. The name clearly distinguishes them from the circuits that generate sinusoids by way of the second approach. In these circuits, a sine wave is obtained by appropriately shaping a triangular waveform. We study waveform-shaping circuits in Section 18.8, following the study of triangular-waveform generators.

Circuits that generate square, triangular, pulse (etc.) waveforms, called **nonlinear oscillators** or **function generators**, employ circuit building blocks known as **multivibrators**. There are three types of multivibrator: the **bistable** (Section 18.4), the **astable** (Section 18.5), and the **monostable** (Section 18.6). The multivibrator circuits presented in this chapter employ op amps and are intended for precision analog applications. Bistable and monostable multivibrator circuits using digital logic gates were studied in Chapter 16.

A general and versatile scheme for the generation of square and triangular waveforms is obtained by connecting a bistable multivibrator and an op-amp integrator in a feedback loop (Section 18.5). Similar results can be obtained using a commercially available versatile IC chip, the 555 timer (Section 18.7).

18.1 Basic Principles of Sinusoidal Oscillators

In this section, we study the basic principles of the design of linear sine-wave oscillators. In spite of the name *linear oscillator,* some form of nonlinearity has to be employed to provide control of the amplitude of the output sine wave. In fact, all oscillators are essentially nonlinear circuits. This complicates the task of analysis and design of oscillators: No longer is one able to apply transform (s-plane) methods directly. Nevertheless, techniques have been developed by which the design of sinusoidal oscillators can be performed in two steps: The first step is a linear one, and frequency-domain methods of feedback circuit analysis can be readily employed. Subsequently, in step 2, a nonlinear mechanism for amplitude control can be provided.

18.1.1 The Oscillator Feedback Loop

The basic structure of a sinusoidal oscillator consists of an amplifier and a frequency-selective network connected in a **positive-feedback loop**, such as that shown in block diagram form in Fig. 18.1. Although no input signal will be present in an actual oscillator circuit, we include an input signal here to help explain the principle of operation. It is important to note that unlike the negative-feedback loop of Fig. 11.1, here the feedback signal x_f is summed with a *positive* sign. Thus the gain-with-feedback is given by

$$A_f(s) = \frac{A(s)}{1 - A(s)\beta(s)} \tag{18.1}$$

where we note the negative sign in the denominator. The loop gain $L(s)$ is given by

$$L(s) \equiv A(s)\beta(s) \tag{18.2}$$

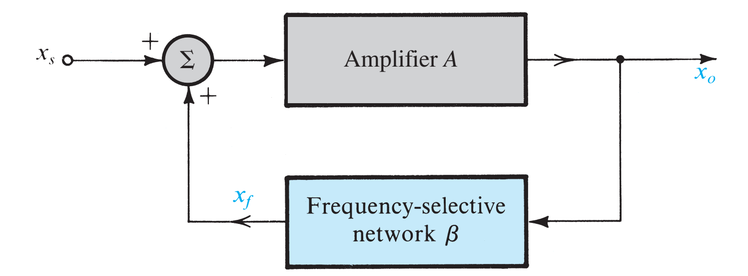


Figure 18.1 The basic structure of a sinusoidal oscillator. A positive-feedback loop is formed by an amplifier and a frequency-selective network. In an actual oscillator circuit, no input signal will be present; here an input signal x_s is employed to help explain the principle of operation.

and the characteristic equation is

$$1 - L(s) = 0 \tag{18.3}$$

18.1.2 The Oscillation Criterion

If at a specific frequency f_0 the loop gain $A\beta$ is equal to unity, it follows from Eq. (18.1) that A_f will be infinite. That is, at this frequency the circuit will have a finite output for zero input signal. Such a circuit is by definition an oscillator. Thus the condition for the feedback loop of Fig. 18.1 to provide sinusoidal oscillations of frequency ω_0 is

$$L(j\omega_0) \equiv A(j\omega_0)\beta(j\omega_0) = 1 \tag{18.4}$$

That is, *at* ω_0 *the phase of the loop gain should be zero and the magnitude of the loop gain should be unity*. This is known as the **Barkhausen criterion**. Note that for the circuit to oscillate at one frequency, the oscillation criterion should be satisfied only at one frequency (i.e., ω_0); otherwise the resulting waveform will not be a simple sinusoid.

An intuitive feeling for the Barkhausen criterion can be gained by considering once more the feedback loop of Fig. 18.1. For this loop to *produce* and *sustain* an output x_o with no input applied ($x_s = 0$), the feedback signal x_f,

$$x_f = \beta x_o$$

should be sufficiently large that when multiplied by A it produces x_o, that is,

$$Ax_f = x_o$$

that is,

$$A\beta x_o = x_o$$

which results in

$$A\beta = 1$$

It should be noted that the *frequency of oscillation* ω_0 is determined solely by the phase characteristics of the feedback loop; the loop oscillates at the frequency for which the phase is zero (or, equivalently, 360°). It follows that the stability of the frequency of oscillation will be determined by the manner in which the phase $\phi(\omega)$ of the feedback loop varies with frequency. A "steep" function $\phi(\omega)$ will result in a more stable frequency. This can be seen if one imagines a change in phase $\Delta\phi$ due to a change in one of the circuit components. If $d\phi/d\omega$ is large, the resulting change in ω_0 will be small, as illustrated in Fig. 18.2.

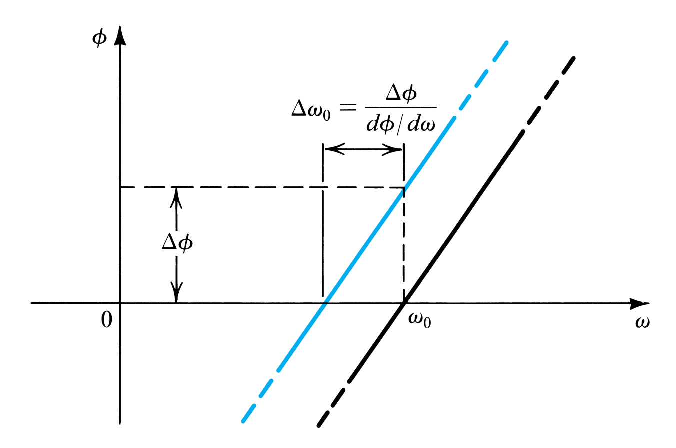


Figure 18.2 Dependence of the oscillator-frequency stability on the slope of the phase response. A steep phase response (i.e., large $d\phi/d\omega$) results in a small $\Delta\omega_0$ for a given change in phase $\Delta\phi$ [resulting from a change (due, for example, to temperature) in a circuit component].

An alternative approach to the study of oscillator circuits consists of examining the circuit poles, which are the roots of the **characteristic equation** (Eq. 18.3). For the circuit to produce **sustained oscillations** at a frequency ω_0 the characteristic equation has to have roots at $s = \pm j\omega_0$. Thus $1 - A(s)\beta(s)$ should have a factor of the form $s^2 + \omega_0^2$.

EXERCISE

18.1 Consider a sinusoidal oscillator formed by connecting an amplifier with a gain of 2 and a second-order bandpass filter in a feedback loop. Find the pole frequency and the center-frequency gain of the filter needed to produce sustained oscillations at 1 kHz.
Ans. 1 kHz; 0.5

18.1.3 Analysis of Oscillator Circuits

Analysis of a given oscillator circuit to determine the frequency of oscillation and the condition for the oscillations to start proceeds in three steps:

1. Break the feedback loop to determine the loop gain $A(s)\beta(s)$. This step is similar to that utilized in Section 11.2 in the analysis of negative-feedback amplifiers.
2. The oscillation frequency ω_0 is found as the frequency for which the phase angle of $A(j\omega)\ \beta(j\omega)$ is zero or, equivalently, 360°.
3. The condition for the oscillations to start is found from

$$|A(j\omega_0)\beta(j\omega_0)| \geq 1$$

Note that making the magnitude of the loop gain slightly greater than unity ensures that oscillations will start.

Example 18.1

Figure 18.3(a) shows a sinusoidal oscillator formed by placing a second-order LCR bandpass filter [see Fig. 17.18(d)] in the feedback path of a positive-gain amplifier. Find the frequency of oscillation ω_0, and the condition for oscillations to start. Assume an ideal op amp.

Solution

To obtain the loop gain, we break the positive-feedback loop at the positive input terminal of the op amp where the input impedance is infinite, apply an input voltage V_i, and find the returned voltage V_r. This results in the circuit in Fig. 18.3(b). Since the op amp is ideal, we can find its output voltage V_o and hence $A(s)$ as

$$A(s) \equiv \frac{V_o}{V_i} = 1 + \frac{r_2}{r_1}$$

The transfer function of the frequency-selective network $\beta(s) \equiv V_r/V_o$ can be found utilizing Eq. (17.39) as

$$\beta(s) \equiv \frac{V_r}{V_o} = \frac{s\dfrac{1}{CR}}{s^2 + s\dfrac{1}{CR} + \dfrac{1}{LC}}$$

While we have found it easy to determine $A(s)$ and $\beta(s)$ separately, our interest is in fact in their product, the loop gain $A(s)\,\beta(s)$,

$$A(s)\beta(s) = \frac{s\dfrac{1}{CR}\left(1 + \dfrac{r_2}{r_1}\right)}{s^2 + s\dfrac{1}{CR} + \dfrac{1}{LC}}$$

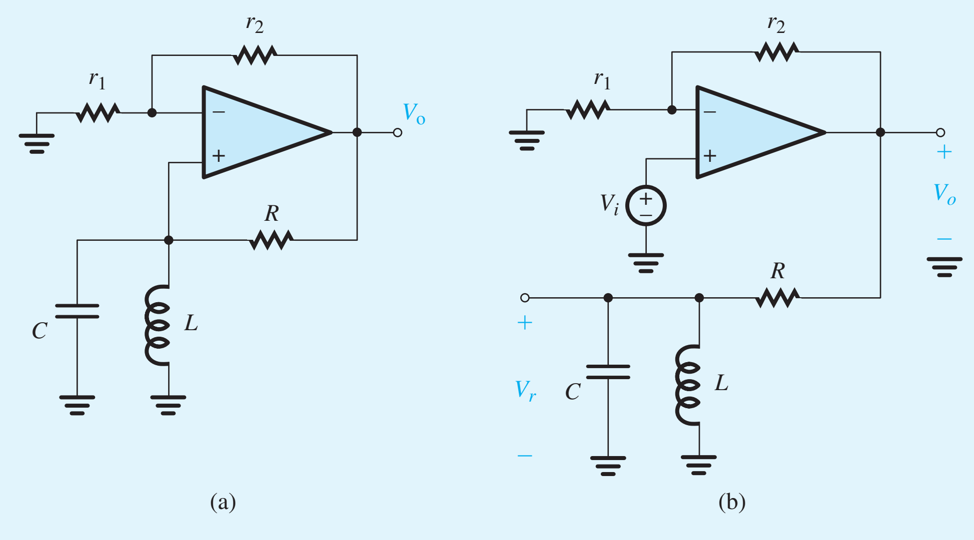


Figure 18.3 **(a)** An oscillator formed by connecting a positive-gain amplifier in a feedback loop with a bandpass RLC circuit. **(b)** Breaking the feedback loop at the input of the op amp to determine $A(s) \equiv V_o(s)/V_i(s)$ and $\beta(s) \equiv V_r(s)/V_o(s)$, and hence the loop gain $A(s)\beta(s)$.

Example 18.1 *continued*

Substituting $s = j\omega$,

$$A(j\omega)\beta(j\omega) = \frac{j\dfrac{\omega}{CR}\left(1 + \dfrac{r_2}{r_1}\right)}{\left(-\omega^2 + \dfrac{1}{LC}\right) + j\dfrac{\omega}{CR}}$$

From this expression we see that the phase angle of $A(j\omega)\ \beta(j\omega)$ will be zero at the value of ω that makes the real part of the denominator zero, thus

$$\omega_0 = 1/\sqrt{LC}$$

At this frequency, the magnitude of the loop gain is given by

$$|A(j\omega_0)\beta(j\omega_0)| = 1 + \frac{r_2}{r_1}$$

Thus oscillation will start for

$$r_2/r_1 \geq 0$$

While $r_2 = 0$ is theoretically sufficient for sustained oscillations, it is usual to design for $r_2/r_1 > 0$ to ensure that oscillations start. This point will be explained in greater detail in the next section.

Before leaving this example, we observe that the results could have been obtained by inspection of the circuit in Fig. 18.3(a). Since the phase angle of the amplifier gain is zero, we examine the LCR bandpass circuit to determine the frequency at which its phase is zero. This in turn is the frequency at which the LC tank has an infinite impedance, which is the resonance frequency $\omega_0 = 1/\sqrt{LC}$. At this frequency the transmission of the bandpass LCR circuit is unity. Thus for oscillation to start, the amplifier gain must be greater than or equal to unity.

An Alternative Analysis Approach There is a simple alternative approach for the analysis of oscillator circuits that does not involve breaking the feedback loop. The method proceeds as follows: We assume that the circuit is oscillating and thus has voltage and current signals at the oscillation frequency ω_0. We analyze the circuit in the usual manner and reduce the equations to a single equation in terms of a single voltage or current variable. Since the voltage or current quantity is not zero (because the circuit is assumed to be oscillating), we can divide by the variable and thus eliminate it. The equation can then be manipulated to the form

$$D(s) = 0$$

where $D(s)$ is a polynomial in s. Substituting $s = j\omega$, we then equate the real and imaginary parts of $D(j\omega)$ to zero. One of the resulting equations yields ω_0 and the other provides the condition for sustained oscillation. This method will be utilized on numerous occasions in the next sections.

EXERCISE

18.2 Apply the alternative analysis method to the circuit of Fig. 18.3(a) and thus determine $D(s)$, $Re[D(j\omega)]$, $\text{Im}[D(j\omega)]$, ω_0, and the condition for sustained oscillation.

Ans. $D(s) = s^2 - s\frac{1}{CR}\frac{r_2}{r_1} + \frac{1}{LC}$; $-\omega^2 + \frac{1}{LC}$; $-\frac{\omega}{CR}\frac{r_2}{r_1}$; $1/\sqrt{LC}$; $r_2 = 0$

18.1.4 Nonlinear Amplitude Control

The oscillation condition, the Barkhausen criterion, discussed in Section 18.1.2, guarantees sustained oscillations in a mathematical sense. It is well known, however, that the parameters of any physical system cannot be maintained constant for any length of time. In other words, suppose we work hard to make $|A\beta| = 1$ at $\omega = \omega_0$, and then the temperature changes and $|A\beta|$ becomes slightly less than unity. Obviously, oscillations will cease in this case. Conversely, if $|A\beta|$ exceeds unity, oscillations will grow in amplitude. We therefore need a mechanism for forcing $|A\beta|$ to remain equal to unity *at the desired value of output amplitude*. This task is accomplished by providing a nonlinear circuit for gain control.

Basically, the function of the gain-control mechanism is as follows: First, to ensure that oscillations will start, one designs the circuit such that $|A\beta|$ is slightly greater than unity. This corresponds to designing the circuit so that the poles are in the right half of the s plane. Thus as the power supply is turned on, oscillations will grow in amplitude. When the amplitude reaches the desired level, the nonlinear network comes into action and causes the loop gain to be reduced to exactly unity. In other words, the poles will be "pulled back" to the $j\omega$ axis. This action will cause the circuit to sustain oscillations at this desired amplitude. If, for some reason, the loop gain is reduced below unity, the amplitude of the sine wave will diminish. This will be detected by the nonlinear network, which will cause the loop gain to increase to exactly unity.

As will be seen, there are two basic approaches to the implementation of the nonlinear amplitude-stabilization mechanism. The first approach makes use of a limiter circuit (see Section 4.6). Oscillations are allowed to grow until the amplitude reaches the level to which the limiter is set. When the limiter comes into operation, the amplitude remains constant. Obviously, the limiter should be "soft" to minimize nonlinear distortion. Such distortion, however, is reduced by the filtering action of the frequency-selective network in the feedback loop. In fact, in one of the oscillator circuits studied in Section 18.2, the sine waves are hard limited, and the resulting square waves are applied to a bandpass filter present in the feedback loop. The "purity" of the output sine waves will be a function of the selectivity of this filter. That is, the higher the Q of the filter, the less the harmonic content of the sine-wave output.

The other mechanism for amplitude control utilizes an element whose resistance can be controlled by the amplitude of the output sinusoid. By placing this element in the feedback circuit so that its resistance determines the loop gain, the circuit can be designed to ensure that the loop gain reaches unity at the desired output amplitude. Diodes, or JFETs operated in the triode region,[1] are commonly employed to implement the controlled-resistance element.

[1]We have not studied JFETs in this book. However, the book website includes material on JFETs and JFET circuits.

EXERCISE

18.3 Assume that the oscillator of Fig. 18.3(a) is designed for a loop-gain magnitude greater than unity so that the amplitude of oscillations grows to the point that the op-amp output saturates. Assume that the saturation levels of the op amp are ±2 V and that the signal at the output of the op amp is nearly a square wave. If the bandpass LCR circuit is of sufficiently high selectivity, what is the amplitude of the sine wave across the LC circuit? (*Hint*: The fundamental frequency component in the Fourier series expansion of a square wave of amplitude V has an amplitude $4V/\pi$.)
Ans. 2.55 V

18.1.5 A Popular Limiter Circuit for Amplitude Control

We conclude this section by presenting a limiter circuit that is frequently employed for the amplitude control of op-amp oscillators, as well as in a variety of other applications. The circuit is more precise and versatile than those presented in Chapter 4.

The limiter circuit is shown in Fig. 18.4(a), and its transfer characteristic is depicted in Fig. 18.4(b). To see how the transfer characteristic is obtained, consider first the case of a small (close to zero) input signal v_I and a small output voltage v_O, so that v_A is positive and v_B is negative. It can be easily seen that both diodes D_1 and D_2 will be off. Thus all of the input current v_I/R_1 flows through the feedback resistance R_f, and the output voltage is given by

$$v_O = -\left(R_f/R_1\right) v_I \tag{18.5}$$

This is the linear portion of the limiter transfer characteristic in Fig. 18.4(b). We now can use superposition to find the voltages at nodes A and B in terms of $\pm V$ and v_O as

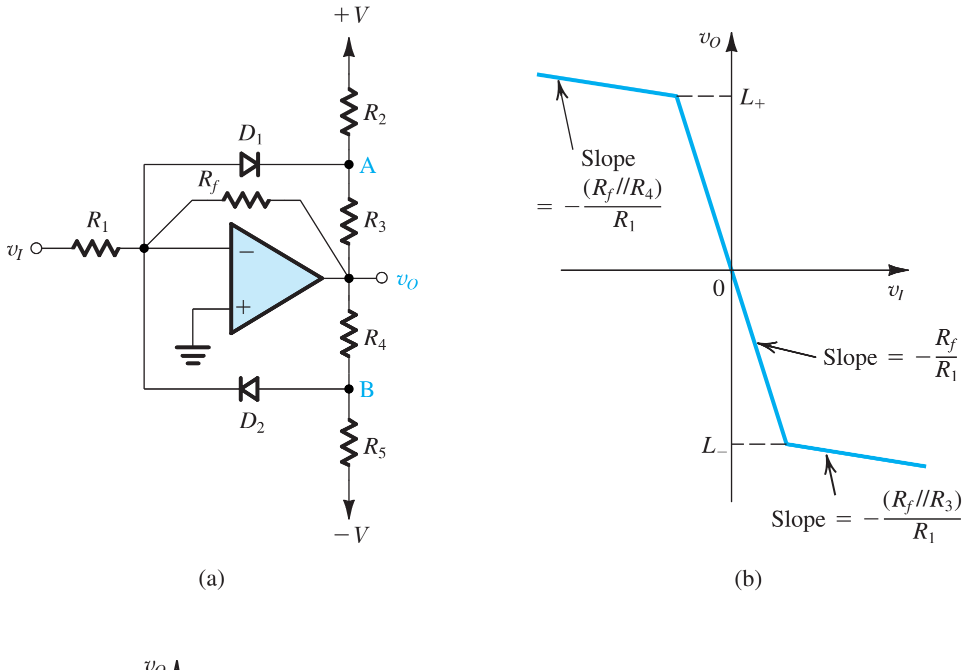
$$v_\mathrm{A} = V\frac{R_3}{R_2+R_3} + v_O\frac{R_2}{R_2+R_3} \tag{18.6}$$

$$v_\mathrm{B} = -V\frac{R_4}{R_4+R_5} + v_O\frac{R_5}{R_4+R_5} \tag{18.7}$$

As v_I goes positive, v_O goes negative (Eq. 18.5), and we see from Eq. (18.7) that v_B will become more negative, thus keeping D_2 off. Equation (18.6) shows, however, that v_A becomes less positive. Then, if we continue to increase v_I, a negative value of v_O will be reached at which v_A becomes −0.7 V or so and diode D_1 conducts. If we use the constant-voltage-drop model for D_1 and denote the voltage drop V_D, the value of v_O at which D_1 conducts can be found from Eq. (18.6). This is the negative limiting level, which we denote L_-,

$$L_- = -V\frac{R_3}{R_2} - V_D\left(1+\frac{R_3}{R_2}\right) \tag{18.8}$$

The corresponding value of v_I can be found by dividing L_- by the limiter gain $-R_f/R_1$. If v_I is increased beyond this value, more current is injected into D_1, and v_A remains at approximately



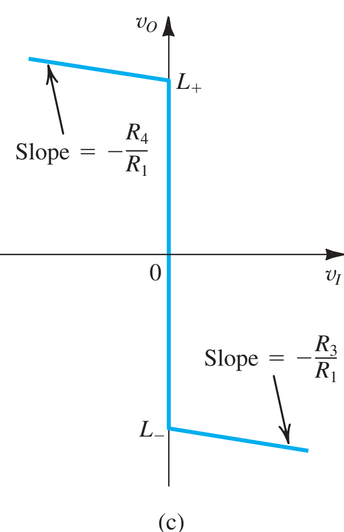


Figure 18.4 **(a)** A popular limiter circuit. **(b)** Transfer characteristic of the limiter circuit; L_- and L_+ are given by Eqs. (18.8) and (18.9), respectively. **(c)** When R_f is removed, the limiter turns into a comparator with the characteristic shown.

$-V_D$. Thus the current through R_2 remains constant, and the additional diode current flows through R_3. Thus R_3 appears in effect in parallel with R_f, and the incremental gain (ignoring the diode resistance) is $-(R_f \parallel R_3)/R_1$. To make the slope of the transfer characteristic small in the limiting region, a low value should be selected for R_3.

The transfer characteristic for negative v_I can be found in a manner identical to that just employed. It can be easily seen that for negative v_I, diode D_2 plays an identical role to that

played by diode D_1 for positive v_I. We can use Eq. (18.7) to find the positive limiting level L_+

$$L_+ = V\frac{R_4}{R_5} + V_D\left(1 + \frac{R_4}{R_5}\right) \tag{18.9}$$

and the slope of the transfer characteristic in the positive limiting region is $-(R_f \| R_4)/R_1$. We thus see that the circuit of Fig. 18.4(a) functions as a soft limiter, with the limiting levels L_+ and L_-, and the limiting gains independently adjustable by the selection of appropriate resistor values.

Finally, we note that increasing R_f results in a higher gain in the linear region while keeping L_+ and L_- unchanged. In the limit, removing R_f altogether results in the transfer characteristic of Fig. 18.4(c), which is that of a comparator. That is, the circuit compares v_I with the comparator reference value of 0 V: $v_I > 0$ results in $v_O \simeq L_-$, and $v_I < 0$ yields $v_O \simeq L_+$.

EXERCISE

18.4 For the circuit of Fig. 18.4(a) with $V = 15$ V, $R_1 = 30$ kΩ, $R_f = 60$ kΩ, $R_2 = R_5 = 9$ kΩ, and $R_3 = R_4 = 3$ kΩ, find the limiting levels and the value of v_I at which the limiting levels are reached. Also determine the limiter gain and the slope of the transfer characteristic in the positive and negative limiting regions. Assume that $V_D = 0.7$ V.
Ans. ±5.93 V; ±2.97 V; –2; –0.095

18.2 Op Amp–RC Oscillator Circuits

In this section we shall study some practical oscillator circuits utilizing op amps and RC networks. These circuits are usually assembled on printed-circuit boards; their frequency of operation extends from very low frequencies to at most 1 MHz.

18.2.1 The Wien-Bridge Oscillator

One of the simplest oscillator circuits is based on the Wien bridge. Figure 18.5 shows a Wien-bridge oscillator without the nonlinear gain-control network. The circuit consists of an op amp connected in the noninverting configuration, with a closed-loop gain of $1 + R_2/R_1$. In the feedback path of this positive-gain amplifier, an RC network is connected. The loop gain can be easily obtained by multiplying the transfer function $V_a(s)/V_o(s)$ of the feedback network by the amplifier gain,

$$L(s) = \left[1 + \frac{R_2}{R_1}\right]\frac{Z_p}{Z_p + Z_s}$$

$$= \frac{1 + R_2/R_1}{1 + Z_sY_p}$$

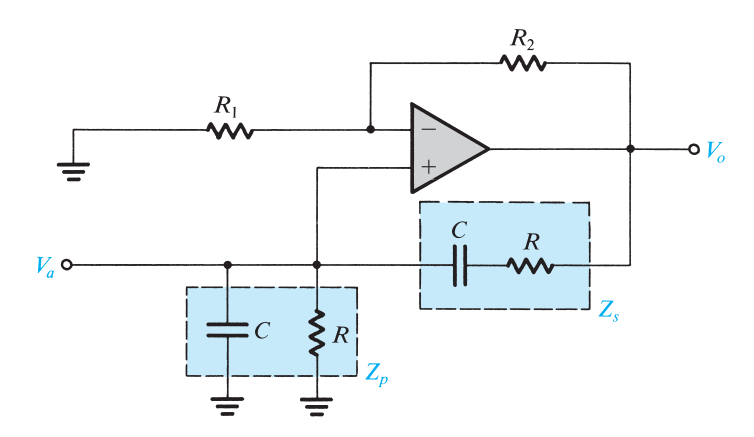


Figure 18.5 A Wien-bridge oscillator without amplitude stabilization.

Thus,

$$L(s) = \frac{1 + R_2/R_1}{3 + sCR + 1/sCR} \tag{18.10}$$

Substituting $s = j\omega$ results in

$$L(j\omega) = \frac{1 + R_2/R_1}{3 + j(\omega CR - 1/\omega CR)} \tag{18.11}$$

The loop gain will be a real number (i.e., the phase will be zero) at one frequency given by

$$\omega_0 CR = \frac{1}{\omega_0 CR}$$

That is,

$$\omega_0 = 1/CR \tag{18.12}$$

Oscillations will start at this frequency if the loop gain is at least unity. This can be achieved by selecting

$$R_2/R_1 = 2 \tag{18.13}$$

To ensure that oscillations will start, one chooses R_2/R_1 slightly greater than 2. The reader can easily verify that if $R_2/R_1 = 2 + \delta$, where δ is a small number, the roots of the characteristic equation $1 - L(s) = 0$ will be in the right half of the s plane.

The amplitude of oscillation can be determined and stabilized by using a nonlinear control network. Two different implementations of the amplitude-controlling function are shown in Figs. 18.6 and 18.7. The circuit in Fig. 18.6 employs a symmetrical feedback limiter of the type studied in Section 18.1.4. It is formed by diodes D_1 and D_2 together with resistors R_3, R_4, R_5, and R_6. The limiter operates in the following manner: At the positive peak of the output voltage v_O, the voltage at node b will exceed the voltage v_1 (which is about $\frac{1}{3}v_O$), and diode D_2 conducts. This will clamp the positive peak to a value determined by R_5, R_6, and the negative power supply. To be specific, the value of the positive output peak can be calculated by setting $v_b = v_1 + V_{D2}$ and writing a node equation at node b while neglecting the current through D_2. Similarly, the negative peak of the output sine wave will be clamped to the value that causes diode D_1 to conduct. The value of the negative peak can be determined by setting $v_a = v_1 - V_{D1}$

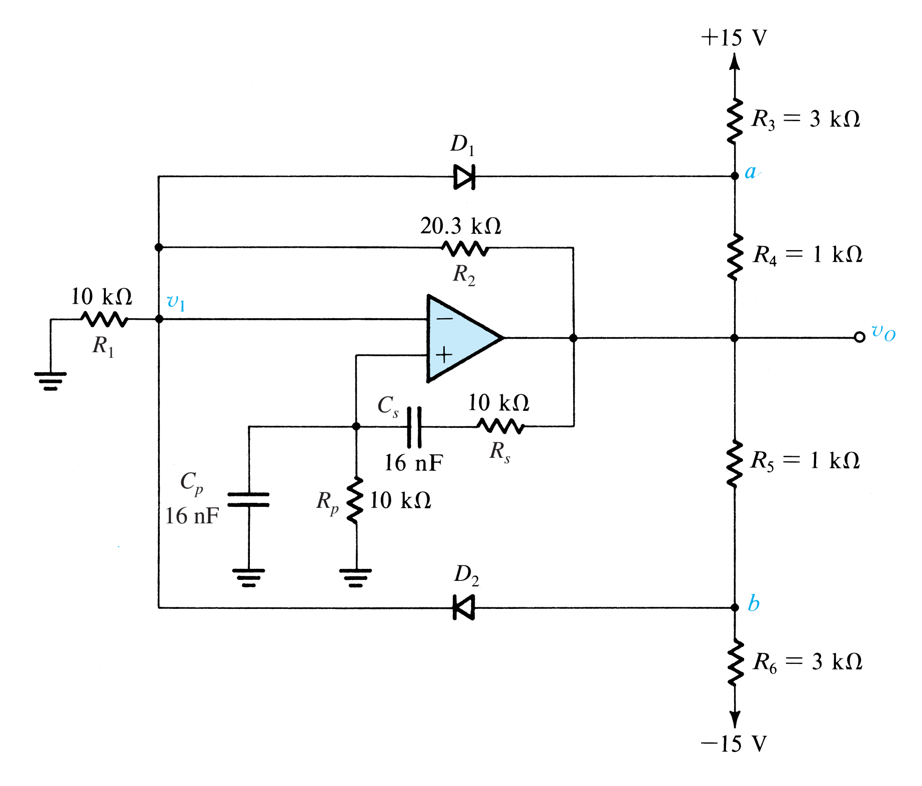


Figure 18.6 A Wien-bridge oscillator with a limiter used for amplitude control.

THE WIEN-BRIDGE OSCILLATOR:

The Wien bridge consisting of four resistors and two capacitors was invented in 1891 by Max Wien, a Prussian physicist, for inductance measurement. Much later, William Hewlett (cofounder in 1939 of Hewlett-Packard), while working toward his master's degree at Stanford University, realized the importance of placing part of the Wien bridge in a positive-feedback loop to form what was called the Wien-bridge oscillator. The first product in 1939 of the new Hewlett-Packard Company was the HP200A, a flexible, precision sine-wave generator using vacuum tubes to implement the amplifier and a tungsten lamp to control the loop gain and thus the amplitude of the sine wave.

EXERCISE

18.5 For the circuit in Fig. 18.6: (a) Disregarding the limiter circuit, find the location of the closed-loop poles. (b) Find the frequency of oscillation. (c) With the limiter in place, find the amplitude of the output sine wave (assume that the diode drop is 0.7 V).

Ans. (a) $(10^5/16)(0.015 \pm j)$; (b) 1 kHz; (c) 21.36 V (peak-to-peak)

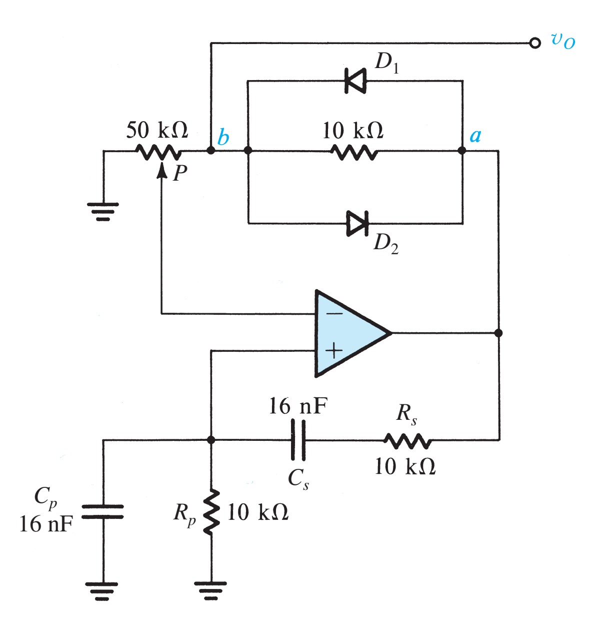


Figure 18.7 A Wien-bridge oscillator with an alternative method for amplitude stabilization.

and writing an equation at node *a* while neglecting the current through D_1. Finally, note that to obtain a symmetrical output waveform, R_3 is chosen equal to R_6, and R_4 equal to R_5.

The circuit of Fig. 18.7 employs an inexpensive implementation of the parameter-variation mechanism of amplitude control. Potentiometer *P* is adjusted until oscillations just start to grow. As the oscillations grow, the diodes start to conduct, causing the effective resistance between *a* and *b* to decrease. Equilibrium will be reached at the output amplitude that causes the loop gain to be exactly unity. The output amplitude can be varied by adjusting potentiometer *P*.

As indicated in Fig. 18.7, the output is taken at point *b* rather than at the op-amp output terminal because the signal at *b* has lower distortion than that at *a*. To appreciate this point, note that the voltage at *b* is proportional to the voltage at the op-amp input terminals and that the latter is a filtered (by the RC network) version of the voltage at node *a*. Node *b*, however, is a high-impedance node, and a buffer will be needed if a load is to be connected.

EXERCISE

18.6 For the circuit in Fig. 18.7, find the following: (a) the setting of potentiometer *P* at which oscillations just start; (b) the frequency of oscillation.
Ans. (a) 20 kΩ to ground; (b) 1 kHz

18.2.2 The Phase-Shift Oscillator

The basic structure of the phase-shift oscillator is shown in Fig. 18.8. It consists of a negative-gain amplifier (–*K*) with a three-section (third-order) RC ladder network in the feedback. The circuit will oscillate at the frequency for which the phase shift of the RC

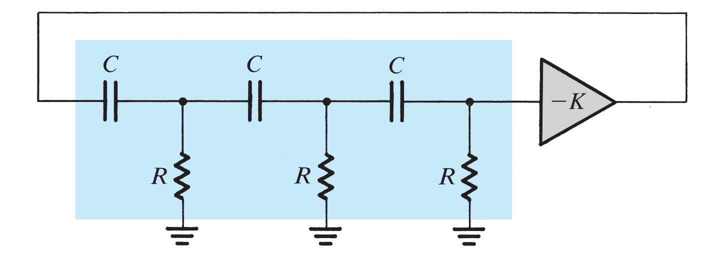


Figure 18.8 A phase-shift oscillator.

network is 180°. Only at this frequency will the total phase shift around the loop be 0° or 360°. Here we should note that the reason for using a three-section RC network is that three is the minimum number of sections (i.e., lowest order) that is capable of producing a 180° phase shift at a finite frequency.

For oscillations to be sustained, the value of K should be equal to the inverse of the magnitude of the RC network transfer function at the frequency of oscillation. However, to ensure that oscillations start, the value of K has to be chosen slightly higher than the value that satisfies the unity-loop-gain condition. Oscillations will then grow in magnitude until limited by some nonlinear control mechanism.

Figure 18.9 shows a practical phase-shift oscillator with a feedback limiter, consisting of diodes D_1 and D_2 and resistors R_1, R_2, R_3, and R_4 for amplitude stabilization. To start oscillations, R_f has to be made slightly greater than the minimum required value. Although the circuit stabilizes more rapidly and provides sine waves with more stable amplitude, if R_f is made much larger than this minimum, the price paid is an increased output distortion.

EXERCISES

18.7 Consider the circuit of Fig. 18.9 *without* the limiter. Break the feedback loop at X and find the loop gain $A\beta \equiv V_o(j\omega)/V_x(j\omega)$ in symbolic form (i.e., do not substitute the numerical values given). To do this, it is easier to start at the output and work backward, finding the various currents and voltages, and eventually V_x in terms of V_o.

Ans. $\dfrac{\omega^2 C^2 R R_f}{4 + j(3\omega CR - 1/\omega CR)}$

18.8 Use the expression derived in Exercise 18.7 to find the frequency of oscillation f_0 and the minimum required value of R_f for oscillations to start in the circuit of Fig. 18.9.

Ans. $\omega_0 = 1/\sqrt{3}CR; R_f \geq 12R; f_0 = 574.3\,\text{Hz}; R_f = 120\text{ k}\Omega$

18.2.3 The Quadrature Oscillator

The **quadrature oscillator** is based on the two-integrator loop studied in Section 17.7. As an active filter, the loop is damped to locate the poles in the left half of the s plane. Here, no such

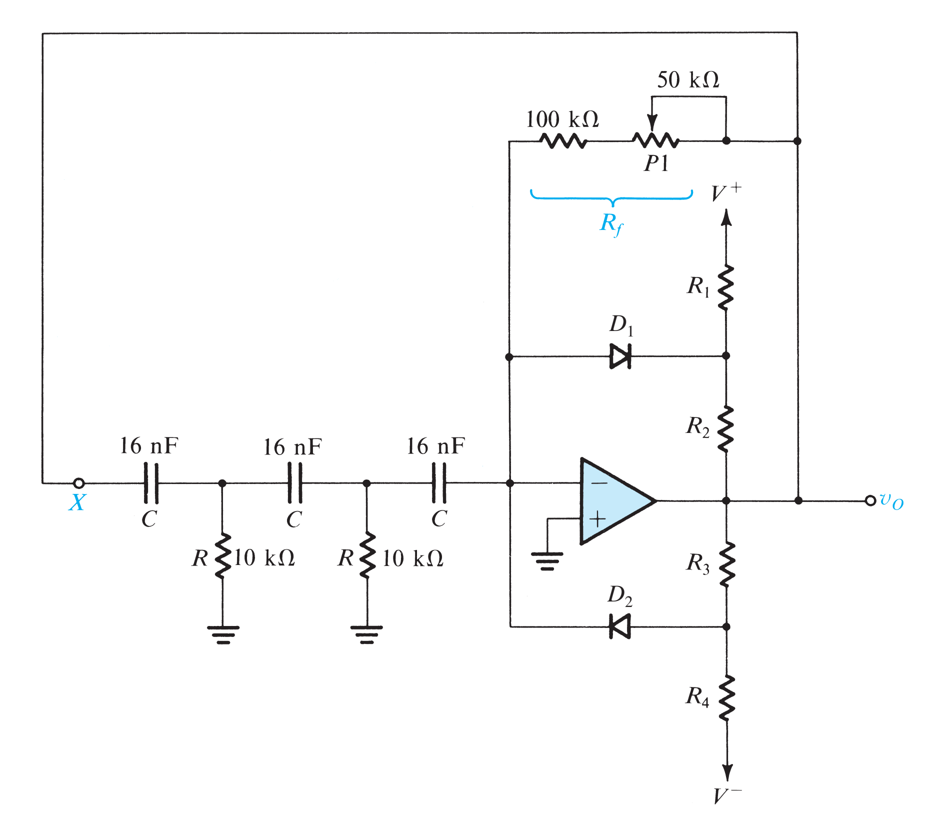


Figure 18.9 A practical phase-shift oscillator with a limiter for amplitude stabilization.

damping will be used, since we wish to locate the poles on the $j\omega$ axis to provide sustained oscillations. In fact, to ensure that oscillations start, the poles are initially located in the right half-plane and then "pulled back" by the nonlinear gain control.

Figure 18.10(a) shows a practical quadrature oscillator. Amplifier 1 is connected as an inverting Miller integrator with a limiter in the feedback for amplitude control. Amplifier 2 is connected as a noninverting integrator [thus replacing the cascade connection of the Miller integrator and the inverter in the two-integrator loop of Fig. 17.25(b)]. To understand the operation of this noninverting integrator, consider the equivalent circuit shown in Fig. 18.10(b). Here, we have replaced the integrator input voltage v_{O1} and the series resistance $2R$ by the Norton equivalent composed of a current source $v_{O1}/2R$ and a parallel resistance $2R$. Now, since $v_{O2} = 2v$, where v is the voltage at the input of op amp 2, the current through R_f will be $(2v - v)/R_f = v/R_f$ in the direction from output to input. Thus R_f gives rise to a negative input resistance, $-R_f$, as indicated in the equivalent circuit of Fig. 18.10(b). Nominally, R_f is made equal to $2R$, and thus $-R_f$ cancels $2R$, and at the input we are left with a current source $v_{O1}/2R$ feeding a capacitor C. The result is that $v = \frac{1}{C}\int_0^t \frac{v_{O1}}{2R} dt$ and $v_{O2} = 2v = \frac{1}{CR}\int_0^t v_{O1} dt$. That is, for $R_f = 2R$, the circuit functions as a perfect noninverting integrator. If, however, R_f is made smaller than $2R$, a net negative resistance appears in parallel with C.

Returning to the oscillator circuit in Fig. 18.10(a), we note that the resistance R_f in the positive-feedback path of op amp 2 is made variable, with a nominal value of $2R$. Decreasing

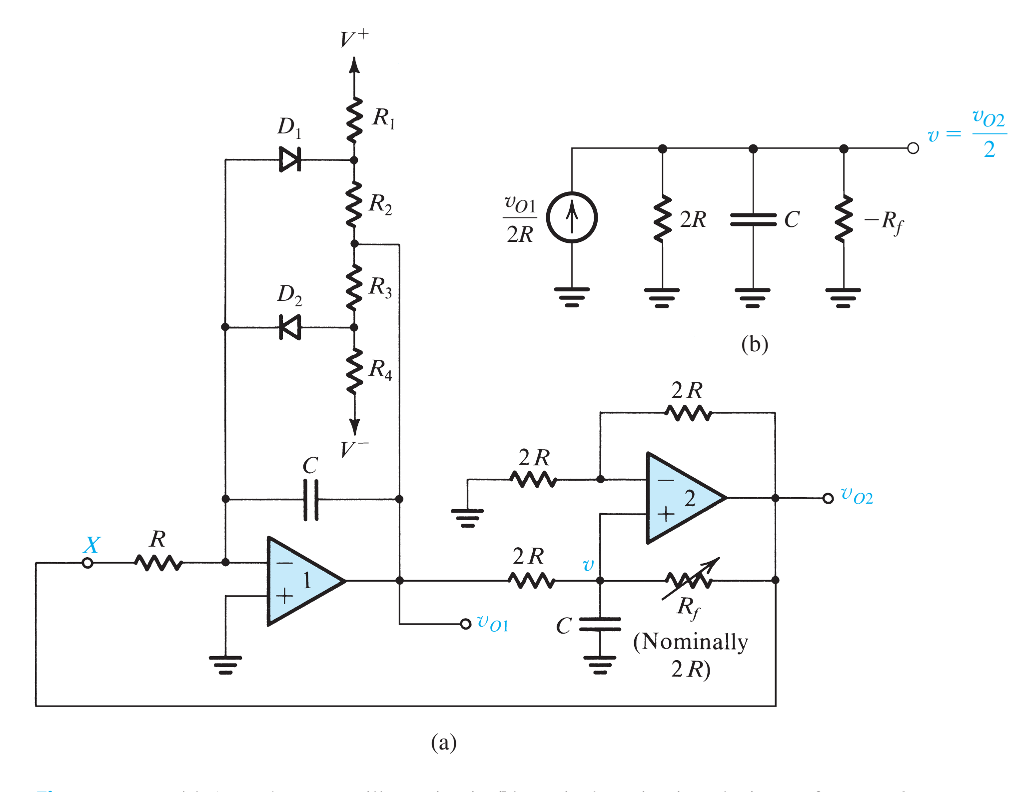


Figure 18.10 **(a)** A quadrature-oscillator circuit. **(b)** Equivalent circuit at the input of op amp 2.

the value of R_f moves the poles to the right half-plane (Problem 18.21) and ensures that the oscillations start. Too much positive feedback, although it results in better amplitude stability, also results in higher output distortion (because the limiter has to operate "harder"). In this regard, note that the output v_{O2} will be "purer" than v_{O1} because of the filtering action provided by the second integrator on the peak-limited output of the first integrator.

If we disregard the limiter and break the loop at X, the loop gain can be obtained as

$$L(s) \equiv \frac{V_{o2}}{V_x} = -\frac{1}{s^2C^2R^2} \tag{18.14}$$

Thus the loop will oscillate at frequency ω_0, given by

$$\omega_0 = \frac{1}{CR} \tag{18.15}$$

Finally, it should be pointed out that the name *quadrature oscillator* is used because the circuit provides two sinusoids with 90° phase difference. This is the case because v_{O2} is the integral of v_{O1}. There are many applications for which quadrature sinusoids are required.

18.2.4 The Active-Filter-Tuned Oscillator

The last oscillator circuit that we shall discuss is quite simple both in principle and in design. Nevertheless, the approach is general and versatile and can result in high-quality (i.e., low-distortion) output sine waves. The basic principle is illustrated in Fig. 18.11. The

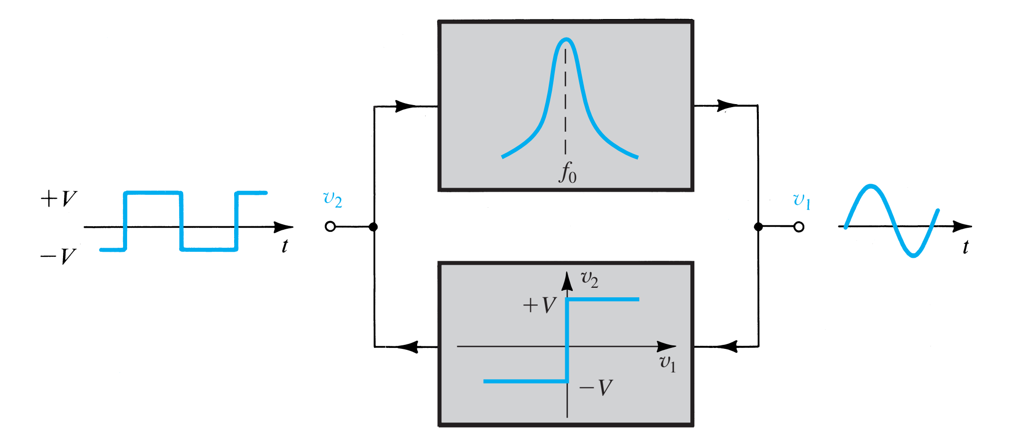


Figure 18.11 Block diagram of the active-filter-tuned oscillator.

circuit consists of a high-Q bandpass filter connected in a positive-feedback loop with a hard limiter. To understand how this circuit works, assume that oscillations have already started. The output of the bandpass filter will be a sine wave whose frequency is equal to the center frequency of the filter, f_0. The sine-wave signal v_1 is fed to the limiter, which produces at its output a square wave whose levels are determined by the limiting levels and whose frequency is f_0. The square wave in turn is fed to the bandpass filter, which filters out the harmonics and provides a sinusoidal output v_1 at the fundamental frequency f_0. Obviously, the purity of the output sine wave will be a direct function of the selectivity (or Q factor) of the bandpass filter.

The simplicity of this approach to oscillator design should be apparent. We have independent control of frequency and amplitude as well as of distortion of the output sinusoid. Any filter circuit with positive gain can be used to implement the bandpass filter. The frequency stability of the oscillator will be directly determined by the frequency stability of the bandpass-filter circuit. Also, a variety of limiter circuits (see Section 4.6) with different degrees of sophistication can be used to implement the limiter block.

Figure 18.12 shows one possible implementation of the active-filter-tuned oscillator. This circuit uses a variation on the bandpass circuit based on the Antoniou inductance-simulation circuit [see Fig. 17.22(c)]. Here resistor R_2 and capacitor C_4 are interchanged. This makes the output of the lower op amp directly proportional to (in fact, twice as large as) the voltage across the resonator, and we can therefore dispense with the buffer amplifier K. The limiter used is a very simple one consisting of a resistance R_1 and two diodes.

EXERCISE

18.9 Using $C = 16\,\text{nF}$, find the value of R such that the circuit of Fig. 18.12 produces 1-kHz sine waves. If the diode drop is 0.7 V, find the peak-to-peak amplitude of the output sine wave. (*Hint:* A square wave with peak-to-peak amplitude of V volts has a fundamental component with $4V/\pi$ volts peak-to-peak amplitude.)
Ans. 10 kΩ; 3.6 V

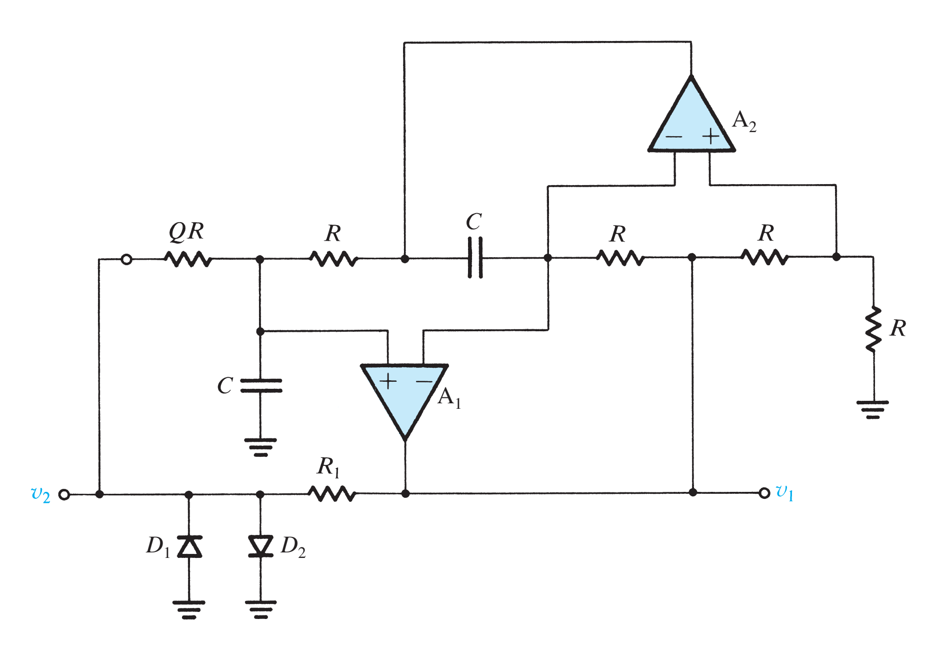


Figure 18.12 A practical implementation of the active-filter-tuned oscillator.

18.2.5 A Final Remark

The op amp–RC oscillator circuits studied are useful for operation in the range 10 Hz to 100 kHz (or perhaps 1 MHz at most). Whereas the lower frequency limit is dictated by the size of passive components required, the upper limit is governed by the frequency-response and slew-rate limitations of op amps. For higher frequencies, circuits that employ transistors together with LC-tuned circuits or crystals are frequently used.[2] These are discussed in Section 18.3.

18.3 LC and Crystal Oscillators

Oscillators utilizing transistors (FETs or BJTs), with LC circuits or crystals as the frequency-selective feedback elements, are used in the frequency range of 100 kHz to hundreds of gigahertz. They exhibit higher Q than the RC types. However, LC oscillators are difficult to tune over wide ranges, and crystal oscillators operate at a single frequency.

18.3.1 The Colpitts and Hartely Oscillators

Figure 18.13 shows two commonly used configurations of LC oscillators. They are known as the **Colpitts oscillator** and the **Hartley oscillator**. Both utilize a parallel LC circuit connected between collector and base (or between drain and gate if a FET is used) with a fraction of the tuned-circuit voltage fed to the emitter (the source in a FET). This feedback is achieved by

[2]Of course, transistors can be used in place of the op amps in the circuits just studied. At higher frequencies, however, better results are obtained with LC-tuned circuits and crystals.

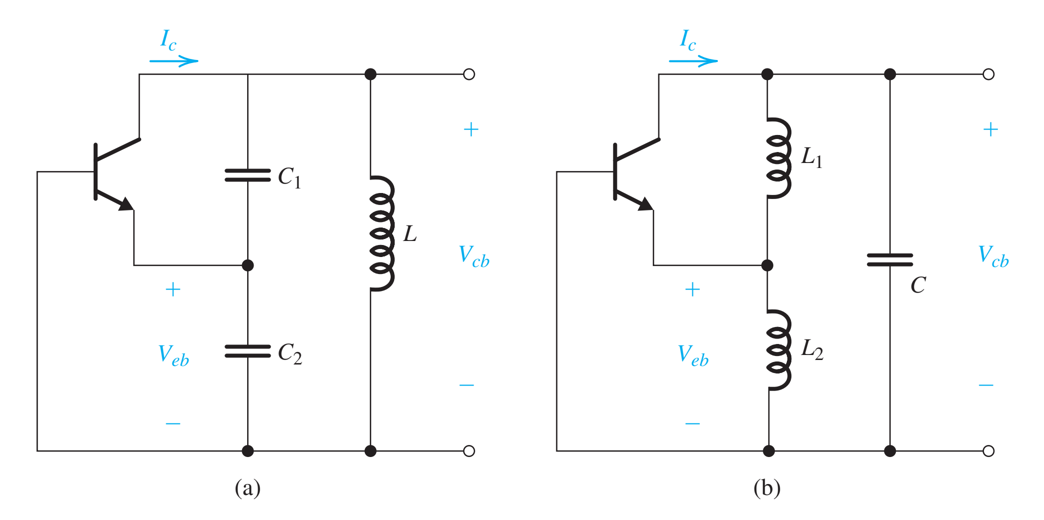


Figure 18.13 Two commonly used configurations of LC-tuned oscillators: **(a)** Colpitts and **(b)** Hartley.

way of a capacitive divider in the Colpitts oscillator and by way of an inductive divider in the Hartley circuit. Observe that in both circuits the voltage V_{eb} gives rise to a current I_c in the direction shown, which in turn results in a positive voltage across the LC circuit. Thus, we do have a positive-feedback loop.

If the frequency of operation is sufficiently low that we can neglect the transistor capacitances, the frequency of oscillation will be determined by the resonance frequency of the parallel-tuned circuit (also known as a *tank circuit* because it behaves as a reservoir for energy storage). Thus for the Colpitts oscillator we have

$$\omega_0 = 1 \bigg/ \sqrt{L\left(\frac{C_1 C_2}{C_1 + C_2}\right)} \tag{18.16}$$

and for the Hartley oscillator we have

$$\omega_0 = 1/\sqrt{(L_1 + L_2)C} \tag{18.17}$$

The ratio L_1/L_2 or C_1/C_2 determines the feedback factor and thus must be adjusted in conjunction with the transistor gain to ensure that oscillations will start. Depending on where the output voltage of the oscillator is to be taken, the appropriate terminal of the transistor can be connected to ground. In this regard, note that any of the three terminals can be grounded without changing the nature of the feedback loop. As an example, we show in Fig. 18.14(a) the Colpitts oscillator with the emitter connected to ground and the output taken at the collector. Although the bias arrangement is not shown, we have included a resistance R that models the combination of the output resistance of the transistor (r_o), the inductor loss, and the input resistance of the circuit to which the oscillator output is connected.

To determine the oscillation condition for the Colpitts oscillator in Fig. 18.14(a), we replace the transistor with its equivalent circuit, as shown in Fig. 18.14(b). To simplify the analysis, we have neglected the transistor capacitance C_μ (C_{gd} for a FET). Capacitance C_π (C_{gs} for a FET), although not shown, can be considered to be a part of C_2. The input resistance r_π (infinite for a FET) has also been neglected, assuming that at the frequency of oscillation $r_\pi \gg (1/\omega C_2)$. Finally, as mentioned earlier, the resistance R includes r_o of the transistor.

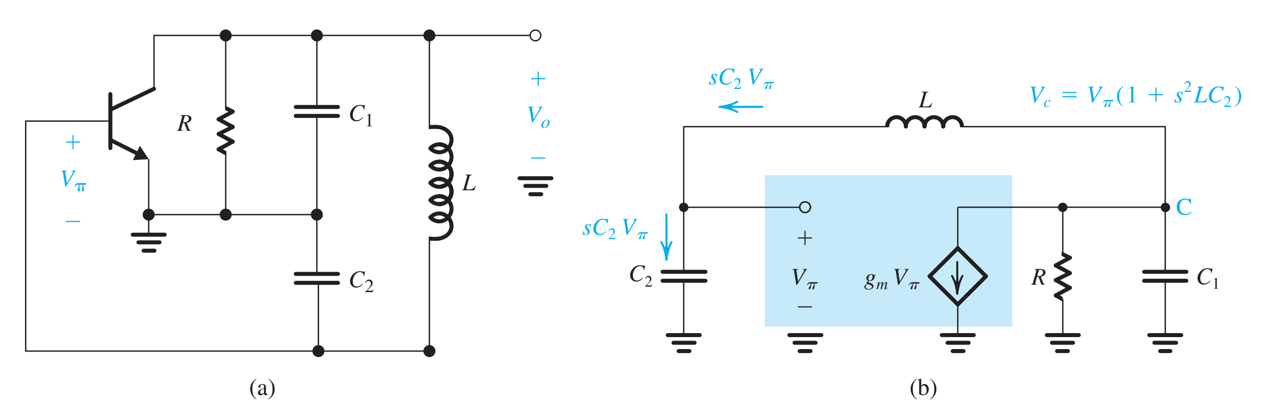


Figure 18.14 (**a**) A Colpitts oscillator in which the emitter is grounded and the output is taken at the collector. (**b**) Equivalent circuit of the Colpitts oscillator of (**a**). To simplify the analysis, C_μ and r_π are neglected. We can consider C_π to be part of C_2, and we can include r_o in R.

Using the alternative analysis approach described in Section 18.1.3, we analyze the circuit as shown in Fig. 18.14(b). A node equation at the transistor collector (node C) yields

$$sC_2V_\pi + g_mV_\pi + \left(\frac{1}{R} + sC_1\right)(1 + s^2LC_2)V_\pi = 0$$

Since $V_\pi \neq 0$ (oscillations have started), it can be eliminated, and the equation can be rearranged in the form

$$s^3LC_1C_2 + s^2(LC_2/R) + s(C_1 + C_2) + \left(g_m + \frac{1}{R}\right) = 0 \quad (18.18)$$

Substituting $s = j\omega$ gives

$$\left(g_m + \frac{1}{R} - \frac{\omega^2LC_2}{R}\right) + j[\omega(C_1 + C_2) - \omega^3LC_1C_2] = 0 \quad (18.19)$$

For oscillations to start, both the real and imaginary parts must be zero. Equating the imaginary part to zero gives the frequency of oscillation as

$$\omega_0 = 1\bigg/\sqrt{L\left(\frac{C_1C_2}{C_1 + C_2}\right)} \quad (18.20)$$

which is the resonance frequency of the tank circuit, as anticipated.[3] Equating the real part to zero and using Eq. (18.20) gives

$$C_2/C_1 = g_mR \quad (18.21)$$

which has a simple physical interpretation: For sustained oscillations, the magnitude of the gain from base to collector (g_mR) must be equal to the inverse of the voltage ratio provided by the capacitive divider, which from Fig. 18.14(a) can be seen to be $V_{eb}/V_{ce} = C_1/C_2$. Of course, for oscillations to start, the loop gain must be made greater than unity, a condition that can be

[3] If r_π is taken into account, the frequency of oscillation can be shown to shift slightly from the value given by Eq. (18.20).

stated in the equivalent form

$$g_m R > C_2/C_1 \qquad (18.22)$$

As oscillations grow in amplitude, the transistor's nonlinear characteristics reduce the effective value of g_m and, correspondingly, reduce the loop gain to unity, thus sustaining the oscillations.

Analysis similar to the foregoing can be carried out for the Hartley circuit (see later: Exercise 18.10). At high frequencies, more accurate transistor models must be used. Alternatively, the y parameters of the transistor can be measured at the intended frequency ω_0, and the analysis can then be carried out using the y-parameter model (see Appendix C). This is usually simpler and more accurate, especially at frequencies above about 30% of the transistor f_T.

As an example of a practical LC oscillator, we show in Fig. 18.15 a discrete-circuit implementation of a Colpitts oscillator, complete with bias details. Here the radio-frequency choke (RFC) provides a high impedance at ω_0 but a low dc resistance.

Finally, a few words are in order on the mechanism that determines the amplitude of oscillations in the LC oscillators discussed above. Unlike the op-amp oscillators that incorporate special amplitude-control circuitry, LC oscillators utilize the nonlinear i_C–v_{BE} characteristics of the BJT (the i_D–v_{GS} characteristics of the FET) for amplitude control. Thus these LC oscillators are known as *self-limiting oscillators*. Specifically, as the oscillations grow in amplitude, the effective gain of the transistor is reduced below its small-signal value. Eventually, an amplitude is reached at which the effective gain is reduced to the point that the Barkhausen criterion is satisfied exactly. The amplitude then remains constant at this value.

Reliance on the nonlinear characteristics of the BJT (or the FET) implies that the collector (drain) current waveform will be nonlinearly distorted. Nevertheless, the output voltage signal will still be a sinusoid of high purity because of the filtering action of the LC circuit. Detailed

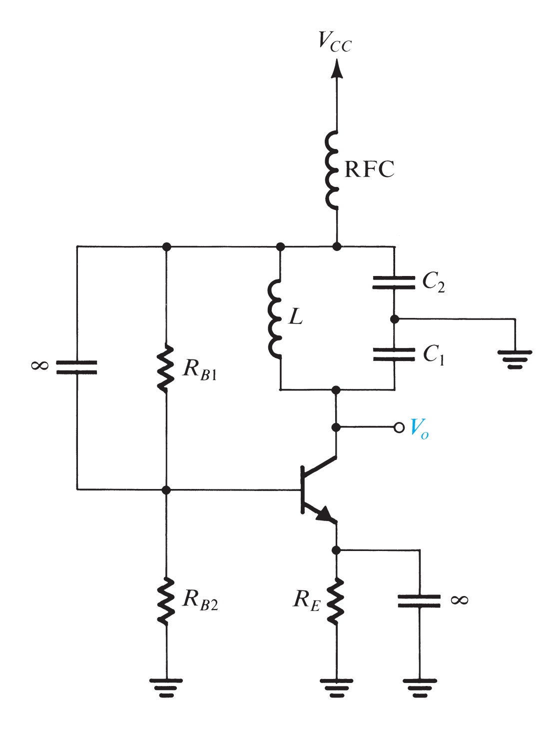


Figure 18.15 Complete discrete-circuit implementation for a Colpitts oscillator.

analysis of amplitude control, which makes use of nonlinear-circuit techniques, is beyond the scope of this book.

EXERCISES

18.10 Show that for the Hartley oscillator of Fig. 18.13(b), the frequency of oscillation is given by Eq. (18.17) and that for oscillations to start $g_m R > (L_1/L_2)$.

D18.11 Using a BJT biased at $I_C = 1$ mA, design a Colpitts oscillator to operate at $\omega_0 = 10^6$ rad/s. Use $C_1 = 0.01\ \mu$F and assume that the coil available has a Q of 100 (this can be represented by a resistance in parallel with C_1 given by $Q/\omega_0 C_1$). Also assume that there is a load resistance at the collector of 2 kΩ and that for the BJT, $r_o = 100$ kΩ. Find C_2 and L.
Ans. 0.66 μF; 100 μH (a somewhat smaller C_2 would be used to allow oscillations to grow in amplitude)

OSCILLATOR PIONEERS:

Heinrich Georg Barkhausen was a German physicist who taught at the Technical University of Dresden. In 1911, at age 29, he became what has since been described as the world's first chaired professor of electrical engineering. Among his many interests were high-frequency oscillators, and he created the enduring Barkhausen stability criterion.

Nevada-born **Ralph Hartley** was a Rhodes Scholar who began work at Western Electric in 1915. He led a transatlantic-radio project for which he invented his tapped-inductor oscillator. As well, he is renowned for having developed a neutralizing scheme to prevent the triode "singing" (self-oscillation) produced by internal input/output capacitance coupling (analogous to that produced by C_μ in a BJT or C_{gd} in a FET).

Edwin Henry Colpitts was a school principal in Newfoundland. He later obtained advanced education in math and physics at Harvard and in 1907 joined Western Electric, where he ultimately led a group in radio telephony. Later, he improved upon the Hartley tapped-inductor oscillator by introducing a more flexible capacitance voltage divider. The circuit now known as the Colpitts oscillator was patented in 1920 as the "oscillation generator." Though invented in the context of vacuum tubes, the Colpitts oscillator is extremely versatile and has marked significance in BJT and MOS designs.

18.3.2 The Cross-Coupled LC Oscillator

A currently popular LC oscillator circuit suitable for fabrication in IC form and capable of operating at frequencies approaching hundreds of gigahertz (for use in wireless transceivers) is shown in Fig. 18.16(a). It consists of a pair of MOSFETs connected in the differential amplifier configuration, each with a parallel LC-tuned circuit load, and with the drain of each connected to the gate of the other. The latter connection gives rise to the "cross-coupled" part of the name.

To see how the cross-coupled LC oscillator works, we first note that, as in a differential amplifier, from a signal point of view each of Q_1 and Q_2 operates in the grounded-source configuration. Next we observe that if we think of Q_1 and Q_2 together with their loads as common-source amplifiers, we see that the cross coupling simply means that the output of each

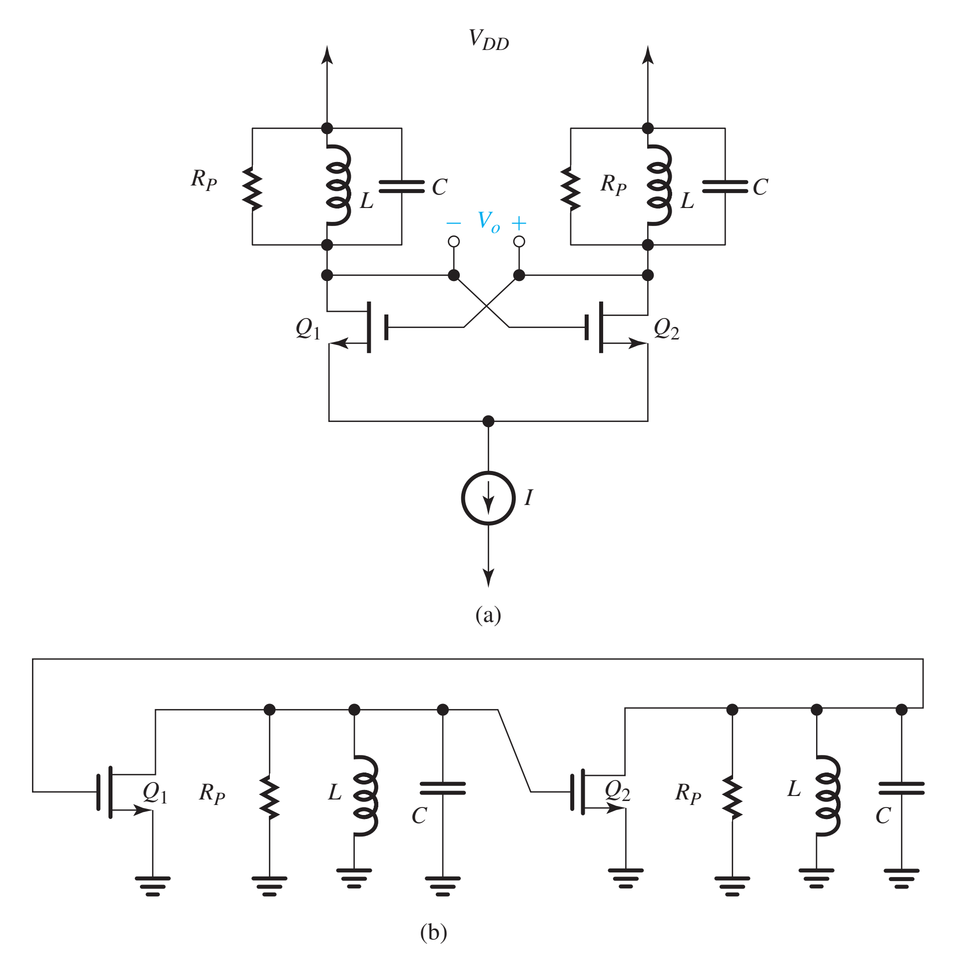


Figure 18.16 **(a)** The cross-coupled LC oscillator. **(b)** Signal equivalent circuit of the cross-coupled oscillator in **(a)**.

amplifier drives the input of the other, resulting in the feedback loop shown in Fig. 18.16(b). Here we have eliminated all dc bias sources to concentrate on the signal operation of the circuit.

Examination of the feedback loop in Fig. 18.16(b) reveals that at the resonance frequency of each of the two tank circuits (i.e., at $\omega = \omega_0 = 1\sqrt{LC}$), the load of each of Q_1 and Q_2 reduces to a resistance $R_p = \omega_0 LQ$, where Q is the quality factor of the inductance. Taking into consideration the output resistance r_o of each of Q_1 and Q_2, we can write for the gain of each of the two stages at $\omega = \omega_0$,

$$A_1 = A_2 = -g_m(R_p \parallel r_o)$$

Thus each stage exhibits a 180° phase shift, for a total phase shift around the loop of 360°. Thus the circuit will provide sustained oscillations at

$$\omega_0 = 1\sqrt{LC} \tag{18.23}$$

provided

$$|A_1 A_2| = [g_m(R_p \parallel r_o)]^2 = 1$$

which reduces to

$$g_m(R_p \parallel r_o) = 1 \tag{18.24}$$

The condition in Eq. (18.24) can be used to determine the minimum required value of g_m at which each of Q_1 and Q_2 is operated for oscillations to be sustained. As usual, to ensure that oscillations start, a somewhat higher value of g_m is used. Amplitude stabilization is provided by the nonlinear MOSFET characteristics.

For applications at very high frequency, the inductor is fabricated on chip by depositing a thin metal film in a spiral shape. Such IC inductors have small values (in the nanohenry range) and, unfortunately, small Q factors as well.

EXERCISE

D18.12 Design the cross-coupled oscillator to operate at $\omega_0 = 10$ Grad/s. The IC inductors available have $L = 10$ nH and $Q = 10$. If the transistor $r_o = 10$ kΩ, find the required value of C and the minimum required value of g_m at which Q_1 and Q_2 are to be operated.
Ans. 1 pF; 1.1 mA/V

18.3.3 Crystal Oscillators

A piezoelectric crystal, such as quartz, exhibits electromechanical-resonance characteristics that are very stable (with time and temperature) and highly selective (having very high Q factors). The circuit symbol of a crystal is shown in Fig. 18.17(a), and its equivalent-circuit model is given in Fig. 18.17(b). The resonance properties are characterized by a large inductance L (as high as hundreds of henrys), a very small series capacitance C_s (as small as 0.0005 pF), a series resistance r representing a Q factor $\omega_0 L/r$ that can be as high as a few hundred thousand, and a parallel capacitance C_p (a few picofarads). Capacitor C_p represents the electrostatic capacitance between the two parallel plates of the crystal. Note that $C_p \gg C_s$.

Since the Q factor is very high, we may neglect the resistance r and express the crystal impedance as

$$Z(s) = 1 \Bigg/ \left[sC_p + \frac{1}{sL + 1/sC_s} \right]$$

which can be manipulated to the form

$$Z(s) = \frac{1}{sC_p} \frac{s^2 + (1/LC_s)}{s^2 + \left[(C_p + C_s)/LC_sC_p\right]} \tag{18.25}$$

From Eq. (18.25) and from Fig. 18.17(b), we see that the crystal has two resonance frequencies: a series resonance at ω_s

$$\omega_s = 1/\sqrt{LC_s} \tag{18.26}$$

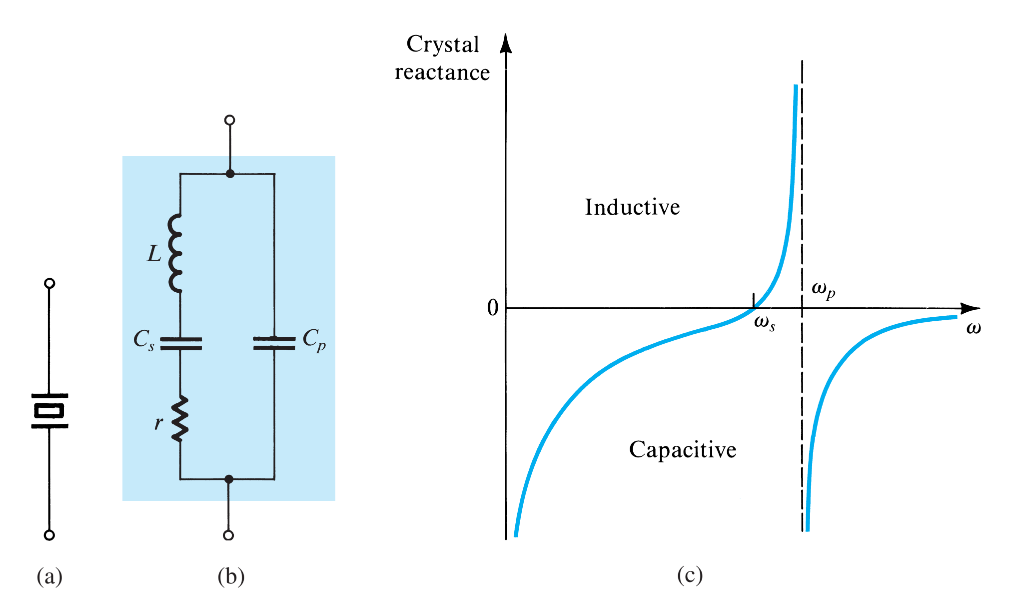


Figure 18.17 A piezoelectric crystal. **(a)** Circuit symbol. **(b)** Equivalent circuit. **(c)** Crystal reactance versus frequency [note that, neglecting the small resistance r, $Z_{\text{crystal}} = jX(\omega)$].

and a parallel resonance at ω_p

$$\omega_p = 1 \Big/ \sqrt{L\left(\frac{C_s C_p}{C_s + C_p}\right)} \tag{18.27}$$

Thus for $s = j\omega$ we can write

$$Z(j\omega) = -j\frac{1}{\omega C_p}\left(\frac{\omega^2 - \omega_s^2}{\omega^2 - \omega_p^2}\right) \tag{18.28}$$

From Eqs. (18.26) and (18.27) we note that $\omega_p > \omega_s$. However, since $C_p \gg C_s$, the two resonance frequencies are very close. Expressing $Z(j\omega) = jX(\omega)$, the crystal reactance $X(\omega)$ will have the shape shown in Fig. 18.17(c). We observe that the crystal reactance is inductive over the very narrow frequency band between ω_s and ω_p. For a given crystal, this frequency band is well defined. Thus we may use the crystal to replace the inductor of the Colpitts oscillator [Fig. 18.13(a)]. The resulting circuit will oscillate at the resonance frequency of the crystal inductance L with the series equivalent of C_s and $(C_p + C_1C_2/(C_1 + C_2))$. Since C_s is much smaller than the three other capacitances, it will be dominant and

$$\omega_0 \simeq 1/\sqrt{LC_s} = \omega_s \tag{18.29}$$

In addition to the basic Colpitts oscillator, a variety of configurations exist for crystal oscillators. Figure 18.18 shows a popular configuration (called the **Pierce oscillator**) utilizing a CMOS inverter (see Section 14.3) as an amplifier. Resistor R_f determines a dc operating point in the high-gain region of the VTC of the CMOS inverter. Resistor R_1 together with capacitor C_1 provides a low-pass filter that discourages the circuit from oscillating at a higher harmonic of the crystal frequency. Note that this circuit also is based on the Colpitts configuration.

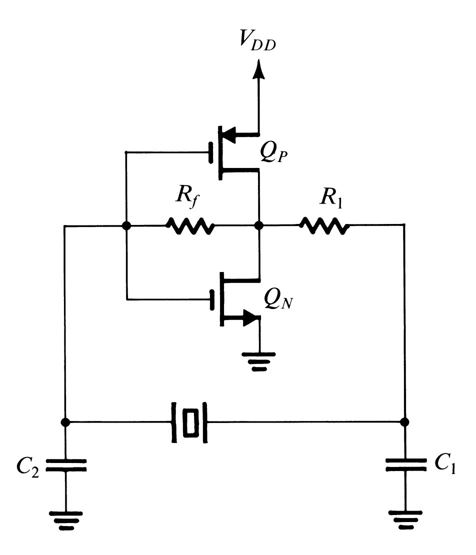


Figure 18.18 A Pierce crystal oscillator utilizing a CMOS inverter as an amplifier.

The extremely stable resonance characteristics and the very high Q factors of quartz crystals result in oscillators with very accurate and stable frequencies. Crystals are available with resonance frequencies in the range of a few kilohertz to hundreds of megahertz. Temperature coefficients of ω_0 of 1 or 2 parts per million (ppm) per degree Celsius are achievable. Unfortunately, however, crystal oscillators, being mechanical resonators, are fixed-frequency circuits.

EXERCISE

18.13 A 2-MHz quartz crystal is specified to have $L = 0.52$ H, $C_s = 0.012$ pF, $C_p = 4$ pF, and $r = 120\ \Omega$. Find f_s, f_p, and Q.
Ans. 2.015 MHz; 2.018 MHz; 55,000

18.4 Bistable Multivibrators

In this section we begin the study of waveform-generating circuits of the other type—nonlinear oscillators or function generators. These devices make use of a special class of circuits known as **multivibrators**. As mentioned earlier, there are three types of multivibrator: bistable, monostable, and astable. This section is concerned with the first, the bistable multivibrator.[4]

As its name indicates, the **bistable multivibrator** has *two stable states*. The circuit can remain in either stable state indefinitely and moves to the other stable state only when appropriately *triggered*.

[4]Digital implementations of multivibrators were presented in Chapter 16. Here, we are interested in implementations utilizing op amps.

18.4.1 The Feedback Loop

Bistability can be obtained by connecting a dc amplifier in a positive-feedback loop having a loop gain greater than unity. Such a feedback loop is shown in Fig. 18.19; it consists of an op amp and a resistive voltage divider in the positive-feedback path. To see how bistability is obtained, consider operation with the positive input terminal of the op amp near ground potential. This is a reasonable starting point, since the circuit has no external excitation. Assume that the electrical noise that is inevitably present in every electronic circuit causes a small positive increment in the voltage v_+. This incremental signal will be amplified by the large open-loop gain A of the op amp, with the result that a much greater signal will appear in the op amp's output voltage v_O. The voltage divider (R_1, R_2) will feed a fraction $\beta \equiv R_1/(R_1+R_2)$ of the output signal back to the positive input terminal of the op amp. If $A\beta$ is greater than unity, as is usually the case, the fed-back signal will be greater than the original increment in v_+. This *regenerative* process continues until eventually the op amp saturates with its output voltage at the positive saturation level, L_+. When this happens, the voltage at the positive input terminal, v_+, becomes $L_+R_1/(R_1+R_2)$, which is positive and thus keeps the op amp in positive saturation. This is one of the two stable states of the circuit.

In the description above we assumed that when v_+ was near zero volts, a positive increment occurred in v_+. Had we assumed the equally probable situation of a negative increment, the op amp would have ended up saturated in the negative direction with $v_O = L_-$ and $v_+ = L_-R_1/(R_1+R_2)$. This is the other stable state.

We thus conclude that the circuit of Fig. 18.19 has two stable states, one with the op amp in positive saturation and the other with the op amp in negative saturation. The circuit can exist in either of these two states indefinitely. We also note that the circuit cannot exist in the state for which $v_+ = 0$ and $v_O = 0$ for any length of time. This is a state of *unstable equilibrium* (also known as a **metastable state**); any disturbance, such as that caused by electrical noise, causes the bistable circuit to switch to one of its two stable states. This is in sharp contrast to the case when the feedback is negative, causing a virtual short circuit to appear between the op amp's input terminals and maintaining this virtual short circuit in the face of disturbances. A physical analogy for the operation of the bistable circuit is depicted in Fig. 18.20.

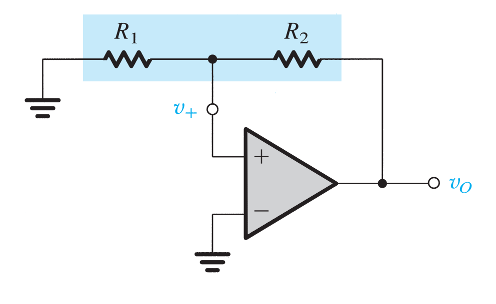


Figure 18.19 A positive-feedback loop capable of bistable operation.

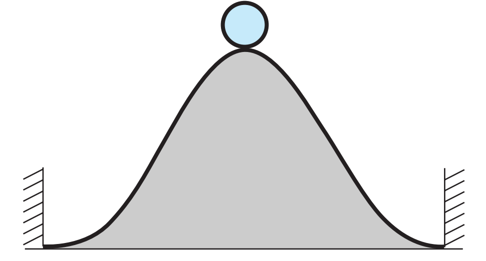


Figure 18.20 A physical analogy for the operation of the bistable circuit. The ball cannot remain at the top of the hill for any length of time (a state of unstable equilibrium or metastability); the inevitably present disturbance will cause the ball to fall to one side or the other, where it can remain indefinitely (the two stable states).

18.4.2 Transfer Characteristic of the Bistable Circuit

The question naturally arises as to how we can make the bistable circuit of Fig. 18.19 change state. To help answer this crucial question, we derive the transfer characteristic of the bistable circuit. Reference to Fig. 18.19 indicates that either of the two nodes that are connected to ground can serve as an input terminal. We investigate both possibilities.

Figure 18.21(a) shows the bistable circuit with a voltage v_I applied to the inverting input terminal of the op amp. To derive the transfer characteristic v_O–v_I, assume that v_O is at one of its two possible levels, say L_+, and thus $v_+ = \beta L_+$. Now as v_I is increased from 0 V, we can see from the circuit that nothing happens until v_I reaches a value equal to v_+ (i.e., βL_+). As

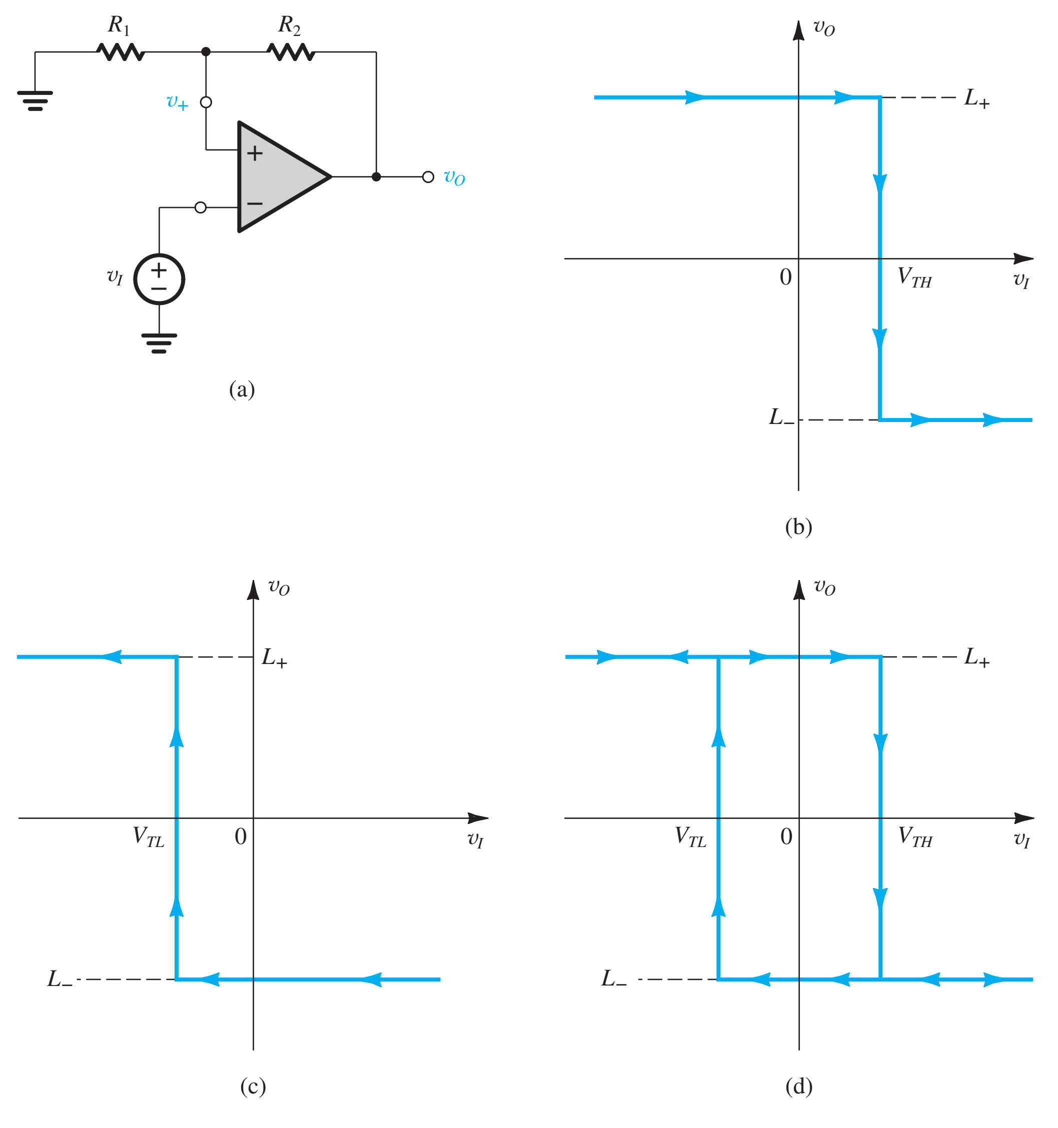


Figure 18.21 **(a)** The bistable circuit of Fig. 18.19 with the negative input terminal of the op amp disconnected from ground and connected to an input signal v_I. **(b)** The transfer characteristic of the circuit in **(a)** for increasing v_I. **(c)** The transfer characteristic for decreasing v_I. **(d)** The complete transfer characteristics.

v_I begins to exceed this value, a net negative voltage develops between the input terminals of the op amp. This voltage is amplified by the open-loop gain of the op amp, and thus v_O goes negative. The voltage divider in turn causes v_+ to go negative, thus increasing the net negative input to the op amp and keeping the regenerative process going. This process culminates in the op amp saturating in the negative direction: that is, with $v_O = L_-$ and, correspondingly, $v_+ = \beta L_-$. It is easy to see that increasing v_I further has no effect on the acquired state of the bistable circuit. Figure 18.21(b) shows the transfer characteristic for increasing v_I. Observe that the characteristic is that of a comparator with a threshold voltage denoted V_{TH}, where $V_{TH} = \beta L_+$.

Next consider what happens as v_I is decreased. Since now $v_+ = \beta L_-$, we see that the circuit remains in the negative-saturation state until v_I goes negative to the point that it equals βL_-. As v_I goes below this value, a net positive voltage appears between the op amp's input terminals. This voltage is amplified by the op-amp gain and thus gives rise to a positive voltage at the op amp's output. The regenerative action of the positive-feedback loop then sets in and causes the circuit eventually to go to its positive saturation state, in which $v_O = L_+$ and $v_+ = \beta L_+$. The transfer characteristic for decreasing v_I is shown in Fig. 18.21(c). Here again we observe that the characteristic is that of a comparator, but with a threshold voltage $V_{TL} = \beta L_-$.

The complete transfer characteristics, v_O–v_I, of the circuit in Fig. 18.21(a) can be obtained by combining the characteristic in Fig. 18.21(b) and (c), as shown in Fig. 18.21(d). As indicated, the circuit changes state at different values of v_I, depending on whether v_I is increasing or decreasing. Thus the circuit is said to exhibit *hysteresis;* the width of the hysteresis is the difference between the high threshold V_{TH} and the low threshold V_{TL}. Also note that the bistable circuit is in effect a comparator with hysteresis. As will be shown shortly, adding hysteresis to a comparator's characteristics can be very beneficial in certain applications. Finally, observe that because the bistable circuit of Fig. 18.21 switches from the positive state ($v_O = L_+$) to the negative state ($v_O = L_-$) as v_I is increased past the positive threshold V_{TH}, the circuit is said to be *inverting*. A bistable circuit with a *noninverting* transfer characteristic will be presented shortly.

18.4.3 Triggering the Bistable Circuit

Returning now to the question of how to make the bistable circuit change state, we observe from the transfer characteristics of Fig. 18.21(d) that if the circuit is in the L_+ state it can be switched to the L_- state by applying an input v_I of value greater than $V_{TH} \equiv \beta L_+$. Such an input causes a net negative voltage to appear between the input terminals of the op amp, which initiates the regenerative cycle that culminates in the circuit switching to the L_- stable state. Here it is important to note that the input v_I merely initiates or *triggers* regeneration. Thus we can remove v_I with no effect on the regeneration process. In other words, v_I can be simply a pulse of short duration. The input signal v_I is thus referred to as a **trigger signal**, or simply a **trigger**.

The characteristics of Fig. 18.21(d) indicate also that the bistable circuit can be switched to the positive state ($v_O = L_+$) by applying a negative trigger signal v_I of magnitude greater than that of the negative threshold V_{TL}.

18.4.4 The Bistable Circuit as a Memory Element

We observe from Fig. 18.21(d) that for input voltages in the range $V_{TL} < v_I < V_{TH}$, the output can be either L_+ or L_-, *depending on the state that the circuit is already in*. Thus, for this input

range, the output is determined by the *previous* value of the trigger signal (the trigger signal that caused the circuit to be in its current state). Thus the circuit exhibits *memory*. Indeed, the bistable multivibrator is the basic memory element of digital systems, as we have seen in Chapter 16. Finally, note that in analog-circuit applications, such as the ones of concern to us in this chapter, the bistable circuit is also known as a **Schmitt trigger**.

18.4.5 A Bistable Circuit with Noninverting Transfer Characteristic

The basic bistable feedback loop of Fig. 18.19 can be used to derive a circuit with noninverting transfer characteristic by applying the input signal v_I (the trigger signal) to the terminal of R_1 that is connected to ground. The resulting circuit is shown in Fig. 18.22(a). To obtain the transfer characteristic we first employ superposition to the linear circuit formed by R_1 and R_2, thus expressing v_+ in terms of v_I and v_O as

$$v_+ = v_I \frac{R_2}{R_1 + R_2} + v_O \frac{R_1}{R_1 + R_2} \tag{18.30}$$

From this equation we see that if the circuit is in the positive stable state with $v_O = L_+$, positive values for v_I will have no effect. To trigger the circuit into the L_- state, v_I must be made negative and of such a value as to make v_+ decrease below zero. Thus the low-threshold V_{TL} can be found by substituting in Eq. (18.30) $v_O = L_+$, $v_+ = 0$, and $v_I = V_{TL}$. The result is

$$V_{TL} = -L_+(R_1/R_2) \tag{18.31}$$

Similarly, Eq. (18.30) indicates that when the circuit is in the negative-output state ($v_O = L_-$), negative values of v_I will make v_+ more negative with no effect on operation. To initiate the regeneration process that causes the circuit to switch to the positive state, v_+ must be made to

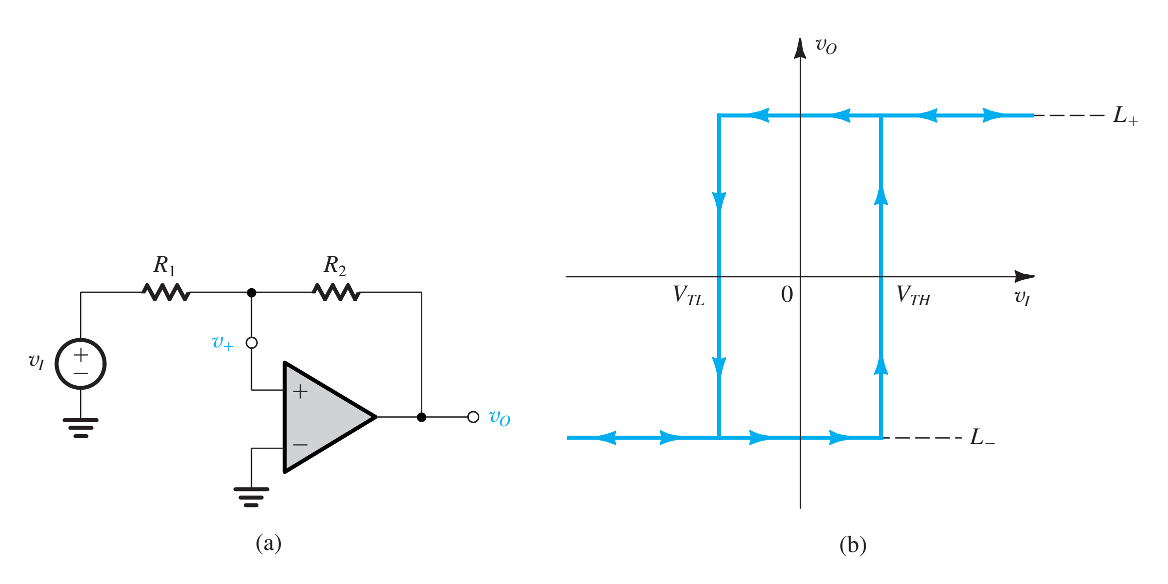


Figure 18.22 **(a)** A bistable circuit derived from the positive-feedback loop of Fig. 18.19 by applying v_I through R_1. **(b)** The transfer characteristic of the circuit in **(a)** is noninverting. [Compare it to the inverting characteristic in Fig. 18.21(d).]

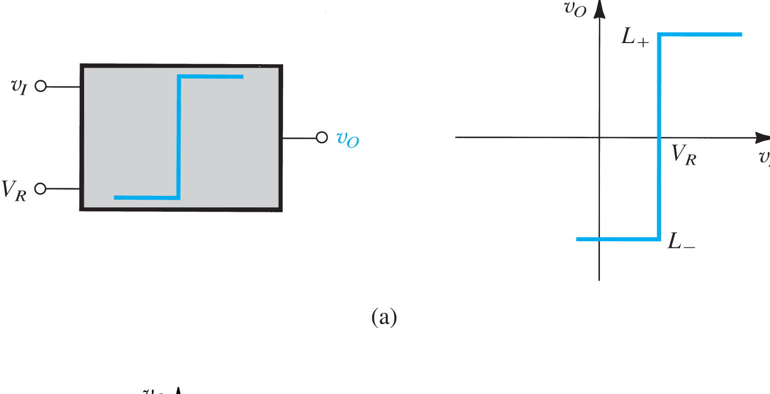
go slightly positive. The value of v_I that causes this to happen is the high-threshold voltage V_{TH}, which can be found by substituting in Eq. (18.30) $v_O = L_-$ and $v_+ = 0$. The result is

$$V_{TH} = -L_-(R_1/R_2) \tag{18.32}$$

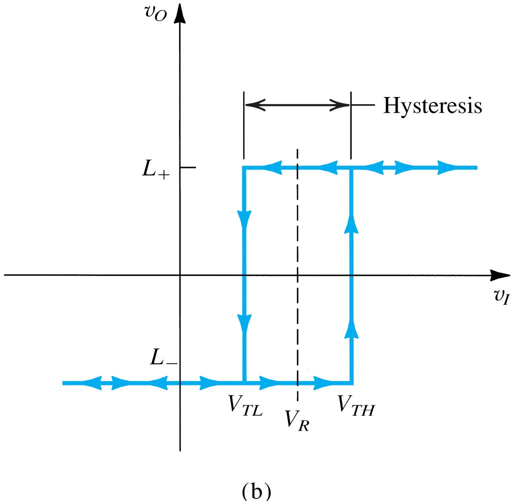
The complete transfer characteristic of the circuit of Fig. 18.22(a) is displayed in Fig. 18.22(b). Observe that a positive triggering signal v_I (of value greater than V_{TH}) causes the circuit to switch to the positive state (v_O goes from L_- to L_+). Thus the transfer characteristic of this circuit is noninverting.

18.4.6 Application of the Bistable Circuit as a Comparator

The comparator is an analog-circuit building block that is used in a variety of applications ranging from detecting the level of an input signal relative to a preset threshold value to the design of analog-to-digital (A/D) converters. Although one normally thinks of the comparator as having a single threshold value [see Fig. 18.23(a)], it is useful in many applications to add hysteresis to the comparator characteristic. If this is done, the comparator



(a)



(b)

Figure 18.23 **(a)** Block diagram representation and transfer characteristic for a comparator having a reference, or threshold, voltage V_R. **(b)** Comparator characteristic with hysteresis.

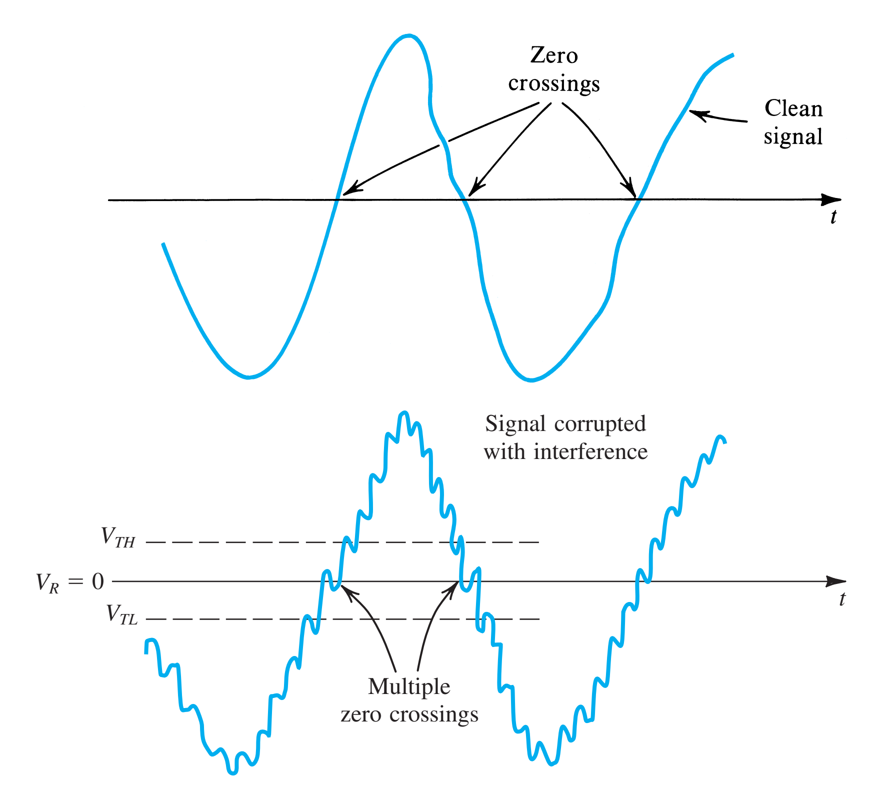


Figure 18.24 Illustrating the use of hysteresis in the comparator characteristic as a means of rejecting interference.

exhibits two threshold values, V_{TL} and V_{TH}, symmetrically placed about the desired reference level, as indicated in Fig. 18.23(b). Usually V_{TH} and V_{TL} are separated by a small amount, say 100 mV.

To demonstrate the need for hysteresis, we consider a common application of comparators. It is required to design a circuit that detects and counts the zero crossings of an arbitrary waveform. Such a function can be implemented using a comparator whose threshold is set to 0 V. The comparator provides a step change at its output every time a zero crossing occurs. Each step change can be used to generate a pulse, and the pulses are fed to a counter circuit.

Imagine now what happens if the signal being processed has—as it usually does have—interference superimposed on it, say of a frequency much higher than that of the signal. It follows that the signal might cross the zero axis a number of times around each of the zero-crossing points we are trying to detect, as shown in Fig. 18.24. The comparator would thus change state a number of times at each of the zero crossings, and our count would obviously be in error. However, if we have an idea of the expected peak-to-peak amplitude of the interference, the problem can be solved by introducing hysteresis of appropriate width in the comparator characteristics. Then, if the input signal is increasing in magnitude, the comparator with hysteresis will remain in the low state until the input level exceeds the high threshold V_{TH}. Subsequently the comparator will remain in the high state even if, owing to interference, the signal decreases below V_{TH}. The comparator will switch to the low state only if the input signal is decreased below the low threshold V_{TL}. The situation is illustrated

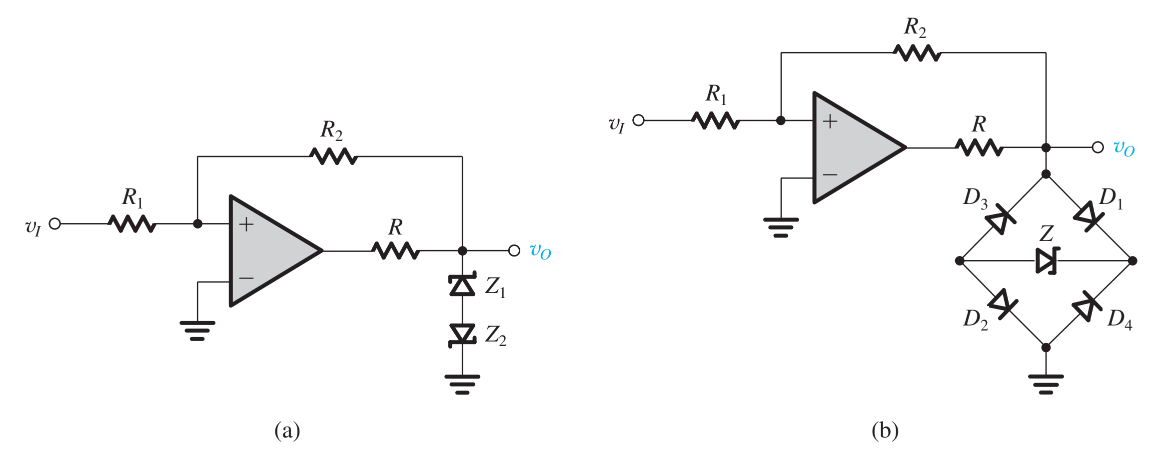


Figure 18.25 Limiter circuits are used to obtain more precise output levels for the bistable circuit. In both circuits the value of R should be chosen to yield the current required for the proper operation of the zener diodes. **(a)** For this circuit $L_+ = V_{Z_1} + V_D$ and $L_- = -(V_{Z_2} + V_D)$, where V_D is the forward diode drop. **(b)** For this circuit $L_+ = V_Z + V_{D_1} + V_{D_2}$ and $L_- = -(V_Z + V_{D_3} + V_{D_4})$.

in Fig. 18.24, from which we see that including hysteresis in the comparator characteristic provides an effective means for rejecting interference (thus providing another form of filtering).

18.4.7 Making the Output Levels More Precise

The output levels of the bistable circuit can be made more precise than the saturation voltages of the op amp are by cascading the op amp with a limiter circuit (see Section 4.6 for a discussion of limiter circuits). Two such arrangements are shown in Fig. 18.25.

EXERCISES

D18.14 The op amp in the bistable circuit of Fig. 18.21(a) has output saturation voltages of ±13 V. Design the circuit to obtain threshold voltages of ±5 V. For $R_1 = 10$ kΩ, find the value required for R_2.
Ans. 16 kΩ

D18.15 If the op amp in the circuit of Fig. 18.22(a) has ±10-V output saturation levels, design the circuit to obtain ±5-V thresholds. Give suitable component values.
Ans. Possible choice: $R_1 = 10$ kΩ and $R_2 = 20$ kΩ

18.16 Consider a bistable circuit with a noninverting transfer characteristic and let $L_+ = -L_- = 10$ V and $V_{TH} = -V_{TL} = 5$ V. If v_I is a triangular wave with a 0-V average, a 10-V peak amplitude, and a 1-ms period, sketch the waveform of v_O. Find the time interval between the zero crossings of v_I and v_O.
Ans. v_O is a square wave with 0-V average, 10-V amplitude, and 1-ms period and is delayed by 125 μs relative to v_I

18.17 Consider an op amp having saturation levels of ±12 V used without feedback, with the inverting input terminal connected to +3 V and the noninverting input terminal connected to v_I. Characterize its operation as a comparator. What are L_+, L_-, and V_R, as defined in Fig. 18.23(a)?
Ans. +12 V; –12 V; +3 V

18.18 In the circuit of Fig. 18.22(a), let $L_+ = -L_- = 10$ V and $R_1 = 1$ kΩ. Find a value for R_2 that gives a hysteresis of 100-mV width.
Ans. 200 kΩ

18.5 Generation of Square and Triangular Waveforms Using Astable Multivibrators

A square waveform can be generated by arranging for a bistable multivibrator to switch states periodically. This can be done by connecting the bistable multivibrator with an RC circuit in a feedback loop, as shown in Fig. 18.26(a). Observe that the bistable multivibrator has an inverting transfer characteristic and can thus be realized using the circuit of Fig. 18.21(a). This results in the circuit of Fig. 18.26(b). We shall show shortly that this circuit has no stable states and thus is appropriately named an **astable multivibrator**.

At this point we wish to remind the reader of an important relationship, which we shall employ on many occasions in the following few sections: A capacitor C that is charging or

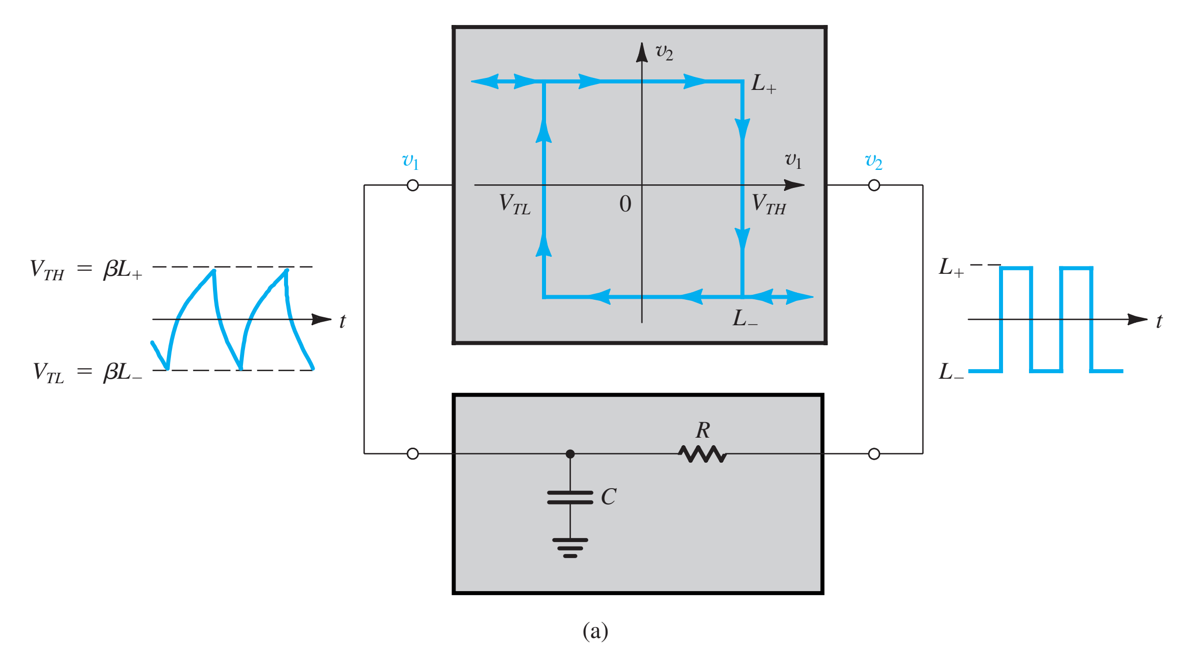


Figure 18.26 **(a)** Connecting a bistable multivibrator with inverting transfer characteristics in a feedback loop with an RC circuit results in a square-wave generator.

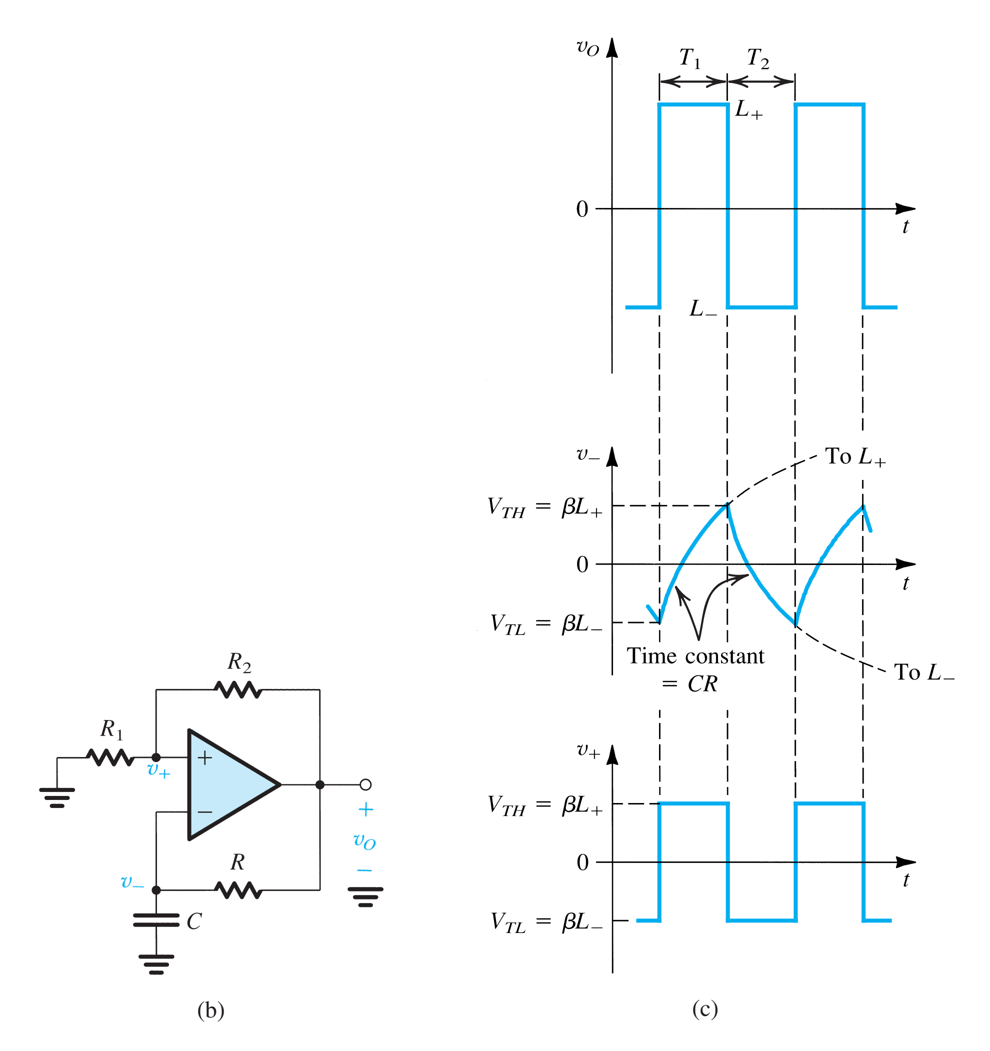


Figure 18.26 *continued* **(b)** The circuit obtained when the bistable multivibrator is implemented with the circuit of Fig. 18.21(a). **(c)** Waveforms at various nodes of the circuit in **(b)**. This circuit is called an astable multivibrator.

discharging through a resistance R toward a final voltage V_∞ has a voltage $v(t)$,

$$v(t) = V_\infty - (V_\infty - V_{0+})e^{-t/\tau}$$

where V_{0+} is the voltage at $t = 0+$ and $\tau = CR$ is the time constant.

18.5.1 Operation of the Astable Multivibrator

To see how the astable multivibrator operates, refer to Fig. 18.26(b) and let the output of the bistable multivibrator be at one of its two possible levels, say L_+. Capacitor C will charge toward this level through resistor R. Thus the voltage across C, which is applied to the negative input terminal of the op amp and thus is denoted v_-, will rise exponentially toward L_+ with a time constant $\tau = CR$. Meanwhile, the voltage at the positive input terminal

of the op amp is $v_+ = \beta L_+$. This situation will continue until the capacitor voltage reaches the positive threshold $V_{TH} = \beta L_+$, at which point the bistable multivibrator will switch to the other stable state, in which $v_O = L_-$ and $v_+ = \beta L_-$. The capacitor will then start discharging, and its voltage, v_-, will decrease exponentially toward L_-. This new state will prevail until v_- reaches the negative threshold $V_{TL} = \beta L_-$, at which time the bistable multivibrator switches to the positive-output state, the capacitor begins to charge, and the cycle repeats itself.

From the preceding description we see that the astable circuit oscillates and produces a square waveform at the output of the op amp. This waveform, and the waveforms at the two input terminals of the op amp, are displayed in Fig. 18.26(c). The period T of the square wave can be found as follows: During the charging interval T_1 the voltage v_- across the capacitor at any time t, with $t = 0$ at the beginning of T_1, is given by (see Appendix E)

$$v_- = L_+ - (L_+ - \beta L_-)e^{-t/\tau}$$

where $\tau = CR$. Substituting $v_- = \beta L_+$ at $t = T_1$ gives

$$T_1 = \tau \ln \frac{1 - \beta(L_-/L_+)}{1 - \beta} \tag{18.33}$$

Similarly, during the discharge interval T_2 the voltage v_- at any time t, with $t = 0$ at the beginning of T_2, is given by

$$v_- = L_- - (L_- - \beta L_+)e^{-t/\tau}$$

Substituting $v_- = \beta L_-$ at $t = T_2$ gives

$$T_2 = \tau \ln \frac{1 - \beta(L_+/L_-)}{1 - \beta} \tag{18.34}$$

Equations (18.33) and (18.34) can be combined to obtain the period $T = T_1 + T_2$. Normally, $L_+ = -L_-$, resulting in symmetrical square waves of period T given by

$$T = 2\tau \ln \frac{1 + \beta}{1 - \beta} \tag{18.35}$$

Note that this square-wave generator can be made to have variable frequency by switching different capacitors C (usually in decades) and by continuously adjusting R (to obtain continuous frequency control within each decade of frequency). Also, the waveform across C can be made almost triangular by using a small value for the parameter β. However, triangular waveforms of superior linearity can be easily generated using the scheme discussed next.

Before leaving this section, however, note that although the astable circuit has no stable states, it has two *quasi-stable* states and remains in each for a time interval determined by the time constant of the RC network and the thresholds of the bistable multivibrator.

EXERCISES

18.19 For the circuit in Fig. 18.26(b), let the op-amp saturation voltages be ±10 V, $R_1 = 100$ kΩ, $R_2 = R = 1$ MΩ, and $C = 0.01$ μF. Find the frequency of oscillation.
Ans. 274 Hz

18.20 Consider a modification of the circuit of Fig. 18.26(b) in which R_1 is replaced by a pair of diodes connected in parallel in opposite directions. For $L_+ = -L_- = 12$ V, $R_2 = R = 10$ kΩ, $C = 0.1$ μF, and the diode voltage as a constant denoted V_D, find an expression for frequency as a function of V_D. If $V_D = 0.70$ V at 25°C with a *TC* of –2 mV/°C, find the frequency at 0°C, 25°C, 50°C, and 100°C. Note that the output of this circuit can be sent to a remotely connected frequency meter to provide a digital readout of temperature.
Ans. $f = 500/\ln[(12 + V_D)/(12 - V_D)]$ Hz; 3995 Hz, 4281 Hz, 4611 Hz, 5451 Hz

18.5.2 Generation of Triangular Waveforms

The exponential waveforms generated in the astable circuit of Fig. 18.24 can be changed to triangular by replacing the low-pass RC circuit with an integrator. (The integrator is, after all, a low-pass circuit with a corner frequency at dc.) The integrator causes linear charging and discharging of the capacitor, thus providing a triangular waveform. The resulting circuit is shown in Fig. 18.27(a). Observe that because the integrator is inverting, it is necessary to invert the characteristics of the bistable circuit. Thus the bistable circuit required here is of the noninverting type and can be implemented using the circuit of Fig. 18.22(a).

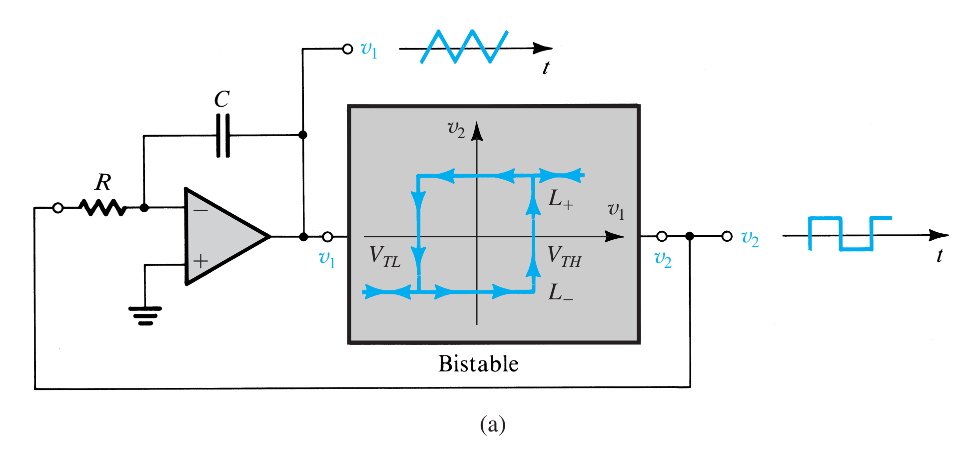
We now proceed to show how the feedback loop of Fig. 18.27(a) oscillates and generates a triangular waveform v_1 at the output of the integrator and a square waveform v_2 at the output of the bistable circuit: Let the output of the bistable circuit be at L_+. A current equal to L_+/R will flow into the resistor R and through capacitor C, causing the output of the integrator to *linearly* decrease with a slope of $-L_+/CR$, as shown in Fig. 18.27(c). This will continue until the integrator output reaches the lower threshold V_{TL} of the bistable circuit, at which point the bistable circuit will switch states, its output becoming negative and equal to L_-. At this moment the current through R and C will reverse direction, and its value will become equal to $|L_-|/R$. It follows that the integrator output will start to increase linearly with a positive slope equal to $|L_-|/CR$. This will continue until the integrator output voltage reaches the positive threshold of the bistable circuit, V_{TH}. At this point the bistable circuit switches, its output becomes positive (L_+), the current into the integrator reverses direction, and the output of the integrator starts to decrease linearly, beginning a new cycle.

From the discussion above, it is relatively easy to derive an expression for the period T of the square and triangular waveforms. During the interval T_1 we have, from Fig. 18.27(c),

$$\frac{V_{TH} - V_{TL}}{T_1} = \frac{L_+}{CR}$$

from which we obtain

$$T_1 = CR\frac{V_{TH} - V_{TL}}{L_+} \tag{18.36}$$



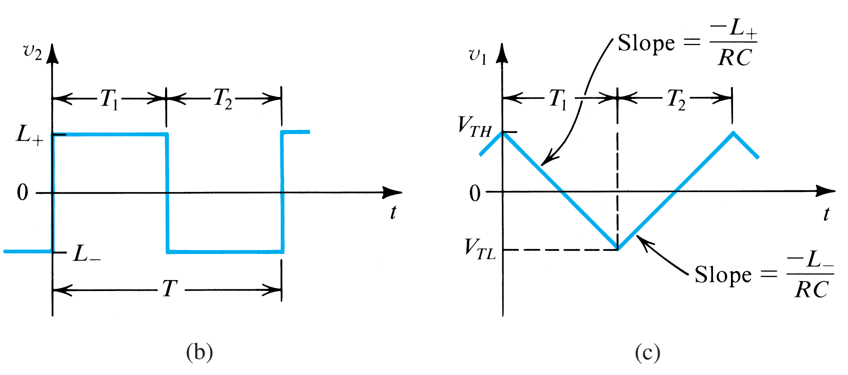
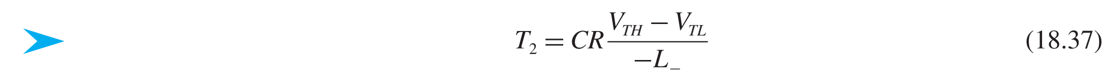


Figure 18.27 A general scheme for generating triangular and square waveforms.

Similarly, during T_2 we have

$$\frac{V_{TH} - V_{TL}}{T_2} = \frac{-L_-}{CR}$$

from which we obtain



$$T_2 = CR\frac{V_{TH} - V_{TL}}{-L_-} \tag{18.37}$$

Thus to obtain symmetrical square waves we design the bistable circuit to have $L_+ = -L_-$.

EXERCISE

D18.21 Consider the circuit of Fig. 18.27(a) with the bistable circuit realized by the circuit in Fig. 18.22(a). If the op amps have saturation voltages of ±10 V, and if a capacitor $C = 0.01$ μF and a resistor $R_1 = 10$ kΩ are used, find the values of R and R_2 (note that R_1 and R_2 are associated with the bistable circuit of Fig. 18.22a) such that the frequency of oscillation is 1 kHz and the triangular waveform has a 10-V peak-to-peak amplitude.
Ans. 50 kΩ; 20 kΩ

18.6 Generation of a Standardized Pulse: The Monostable Multivibrator

In some applications the need arises for a pulse of known height and width generated in response to a trigger signal. Because the width of the pulse is predictable, its trailing edge can be used for timing purposes—that is, to initiate a particular task at a specified time. Such a standardized pulse can be generated by the third type of multivibrator, the **monostable multivibrator**.

The monostable multivibrator has one stable state in which it can remain indefinitely. It also has a quasi-stable state to which it can be triggered and in which it stays for a predetermined interval equal to the desired width of the output pulse. When this interval expires, the monostable multivibrator returns to its stable state and remains there, awaiting another triggering signal. The action of the monostable multivibrator has given rise to its alternative name, the *one-shot*.

Figure 18.28(a) shows an op-amp monostable circuit. We observe that this circuit is an augmented form of the astable circuit of Fig. 18.26(b). Specifically, a clamping diode D_1 is added across the capacitor C_1, and a trigger circuit composed of capacitor C_2, resistor R_4, and diode D_2 is connected to the noninverting input terminal of the op amp. The circuit operates as follows: In the stable state, which prevails in the absence of the triggering signal, the output of the op amp is at L_+ and diode D_1 is conducting through R_3 and thus clamping the voltage v_B to one diode drop above ground. We select R_4 much larger than R_1, so that diode D_2 will be conducting a very small current and the voltage v_C will be very closely determined by the

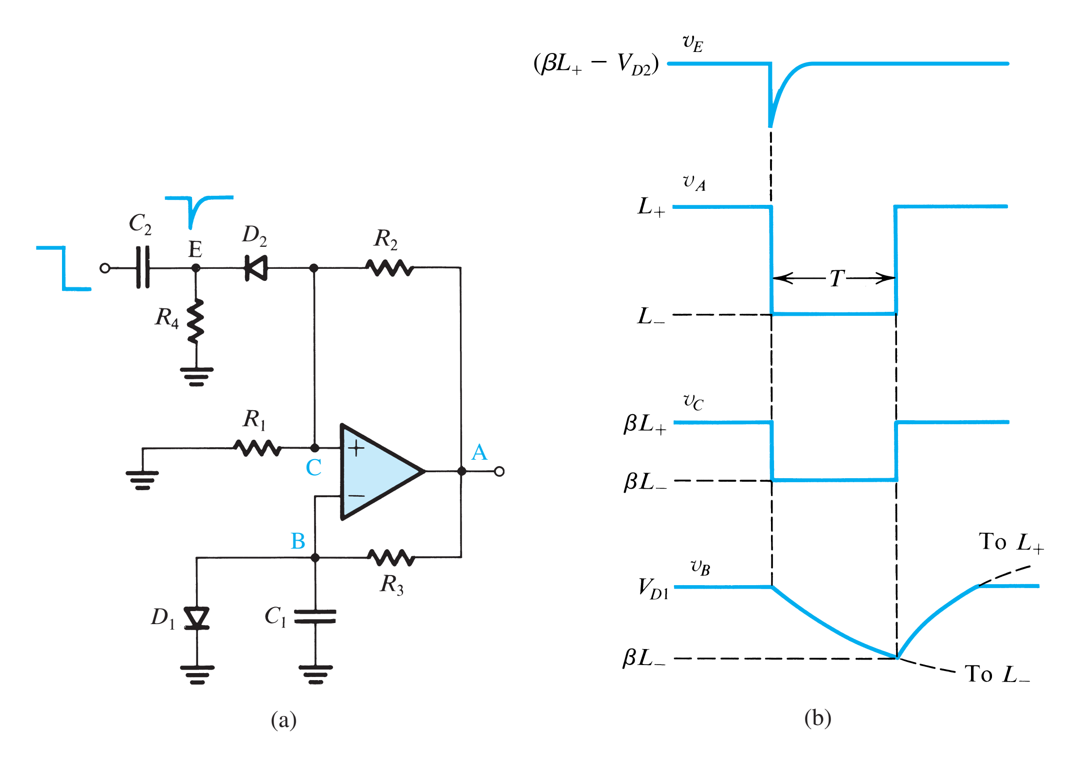


Figure 18.28 **(a)** An op-amp monostable circuit. **(b)** Signal waveforms in the circuit of **(a)**.

voltage divider R_1, R_2. Thus $v_C = \beta L_+$, where $\beta = R_1/(R_1 + R_2)$. The stable state is maintained because βL_+ is greater than V_{D1}.

Now consider the application of a negative-going step at the trigger input and refer to the signal waveforms shown in Fig. 18.28(b). The negative triggering edge is coupled to the cathode of diode D_2 via capacitor C_2, and thus D_2 conducts heavily and pulls node C down. If the trigger signal is of sufficient height to cause v_C to go below v_B, the op amp will see a net negative input voltage and its output will switch to L_-. This in turn will cause v_C to go negative to βL_-, keeping the op amp in its newly acquired state. Note that D_2 will then cut off, thus isolating the circuit from any further changes at the trigger input terminal.

The negative voltage at A causes D_1 to cut off, and C_1 begins to discharge exponentially toward L_- with a time constant C_1R_3. The monostable multivibrator is now in its *quasi-stable state*, which will prevail until the declining v_B goes below the voltage at node C, which is βL_-. At this instant the op-amp output switches back to L_+ and the voltage at node C goes back to βL_+. Capacitor C_1 then charges toward L_+ until diode D_1 turns on and the circuit returns to its stable state.

From Fig. 18.28(b), we observe that a negative pulse is generated at the output during the quasi-stable state. The duration T of the output pulse is determined from the exponential waveform of v_B,

$$v_B(t) = L_- - (L_- - V_{D1})e^{-t/C_1R_3}$$

by substituting $v_B(T) = \beta L_-$,

$$\beta L_- = L_- - (L_- - V_{D1})e^{-T/C_1R_3}$$

which yields

$$T = C_1R_3 \ln\left(\frac{V_{D1} - L_-}{\beta L_- - L_-}\right) \tag{18.38}$$

For $V_{D1} \ll |L_-|$, this equation can be approximated by



$$T \simeq C_1R_3 \ln\left(\frac{1}{1-\beta}\right) \tag{18.39}$$

Finally, note that the monostable circuit should not be triggered again until capacitor C_1 has been recharged to V_{D1}; otherwise the resulting output pulse will be shorter than normal. This recharging time is known as the **recovery period**. Circuit techniques exist for shortening the recovery period.

EXERCISE

18.22 For the monostable circuit of Fig. 18.28(a), find the value of R_3 that will result in a 100-μs output pulse for $C_1 = 0.1$ μF, $\beta = 0.1$, $V_D = 0.7$ V, and $L_+ = -L_- = 12$ V.
Ans. 6171 Ω

18.7 Integrated-Circuit Timers

Commercially available integrated-circuit packages exist that contain the bulk of the circuitry needed to implement monostable and astable multivibrators with precise characteristics. In this section we discuss the most popular of such ICs, the **555 timer**. Introduced in 1972 by the Signetics Corporation as a bipolar integrated circuit, the 555 is also available in CMOS technology and from a number of manufacturers.[5]

18.7.1 The 555 Circuit

Figure 18.29 shows a block diagram representation of the 555 timer circuit (for the actual circuit, refer to Grebene, 1984). The circuit consists of two comparators, an SR flip-flop, and a transistor Q_1 that operates as a switch. One power supply (V_{CC}) is required for operation, with the supply voltage typically 5 V. A resistive voltage divider, consisting of the three equal-valued resistors labeled R_1, is connected across V_{CC} and establishes the reference (threshold) voltages for the two comparators. These are $V_{TH} = \frac{2}{3}V_{CC}$ for comparator 1 and $V_{TL} = \frac{1}{3}V_{CC}$ for comparator 2.

We studied SR flip-flops in Chapter 16. For our purposes here we note that an SR flip-flop is a bistable circuit having complementary outputs, denoted Q and $\overline{Q}$. In the *set* state, the output at Q is "high" (approximately equal to V_{CC}) and that at $\overline{Q}$ is "low" (approximately equal to 0 V). In the other stable state, termed the *reset* state, the output at Q is low and that at $\overline{Q}$ is

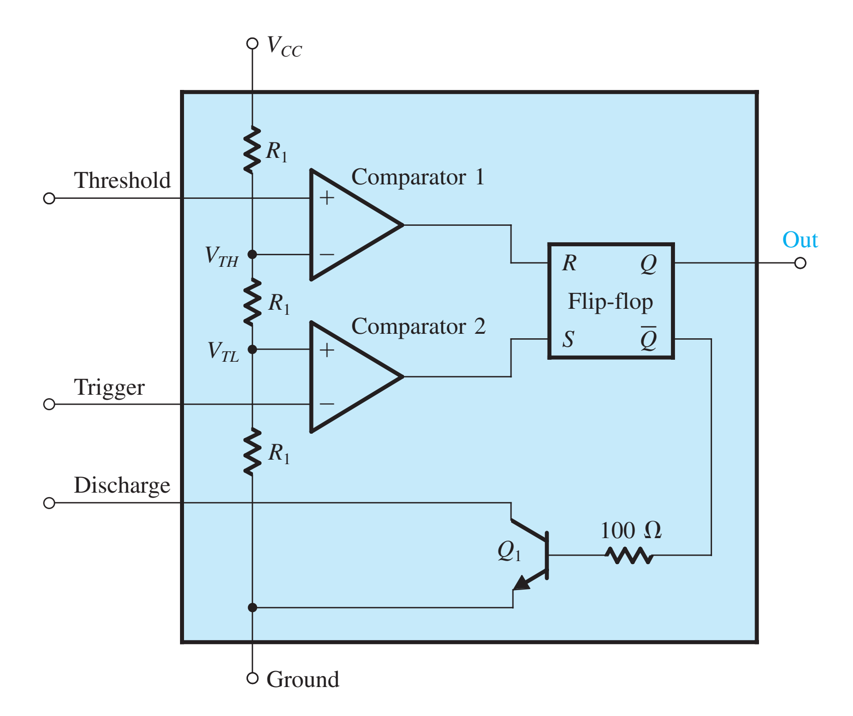


Figure 18.29 A block diagram representation of the internal circuit of the 555 integrated-circuit timer.

[5]In a recent article in *IEEE Spectrum* (May 2009), the 555 was selected as one of the "25 Microchips That Shook the World."

high. The flip-flop is set by applying a high level (V_{CC}) to its set input terminal, labeled S. To reset the flip-flop, a high level is applied to the reset input terminal, labeled R. Note that the reset and set input terminals of the flip-flop in the 555 circuit are connected to the outputs of comparator 1 and comparator 2, respectively.

The positive-input terminal of comparator 1 is brought out to an external terminal of the 555 package, labeled Threshold. Similarly, the negative-input terminal of comparator 2 is connected to an external terminal labeled Trigger, and the collector of transistor Q_1 is connected to a terminal labeled Discharge. Finally, the Q output of the flip-flop is connected to the output terminal of the timer package, labeled Out.

18.7.2 Implementing a Monostable Multivibrator Using the 555 IC

Figure 18.30(a) shows a monostable multivibrator implemented using the 555 IC together with an external resistor R and an external capacitor C. In the stable state the flip-flop will be in the reset state, and thus its $\overline{Q}$ output will be high, turning on transistor Q_1. Transistor Q_1 will be saturated, and thus v_C will be close to 0 V, resulting in a low level at the output of comparator 1. The voltage at the trigger input terminal, labeled v_{trigger}, is kept high (greater than V_{TL}), and thus the output of comparator 2 also will be low. Finally, note that since the flip-flop is in the reset state, Q will be low and thus v_O will be close to 0 V.

To trigger the monostable multivibrator, a negative input pulse is applied to the trigger input terminal. As v_{trigger} goes below V_{TL}, the output of comparator 2 goes to the high level, thus setting the flip-flop. Output Q of the flip-flop goes high, and thus v_O goes high, and output $\overline{Q}$ goes low, turning off transistor Q_1. Capacitor C now begins to charge up through resistor R, and its voltage v_C rises exponentially toward V_{CC}, as shown in Fig. 18.30(b). The monostable multivibrator is now in its quasi-stable state. This state prevails until v_C reaches and begins to exceed the threshold of comparator 1, V_{TH}, at which time the output of comparator 1 goes high, resetting the flip-flop. Output $\overline{Q}$ of the flip-flop now goes high and turns on transistor Q_1. In turn, transistor Q_1 rapidly discharges capacitor C, causing v_C to go to 0 V. Also, when the flip-flop is reset, its Q output goes low, and thus v_O goes back to 0 V. The monostable multivibrator is now back in its stable state and is ready to receive a new triggering pulse.

From the description above we see that the monostable multivibrator produces an output pulse v_O as indicated in Fig. 18.30(b). The width of the pulse, T, is the time interval that the monostable multivibrator spends in the quasi-stable state; it can be determined by reference to the waveforms in Fig. 18.30(b) as follows: Denoting the instant at which the trigger pulse is applied as $t = 0$, the exponential waveform of v_C can be expressed as

$$v_C = V_{CC}\left(1 - e^{-t/CR}\right) \tag{18.40}$$

Substituting $v_C = V_{TH} = \frac{2}{3}V_{CC}$ at $t = T$ gives

$$T = CR\ln 3 \simeq 1.1CR \tag{18.41}$$

Thus the pulse width is determined by the external components C and R, which can be selected to have values as precise as desired.

18.7.3 An Astable Multivibrator Using the 555 IC

Figure 18.31(a) shows the circuit of an astable multivibrator employing a 555 IC, two external resistors, R_A and R_B, and an external capacitor C. To see how the circuit operates, refer to the

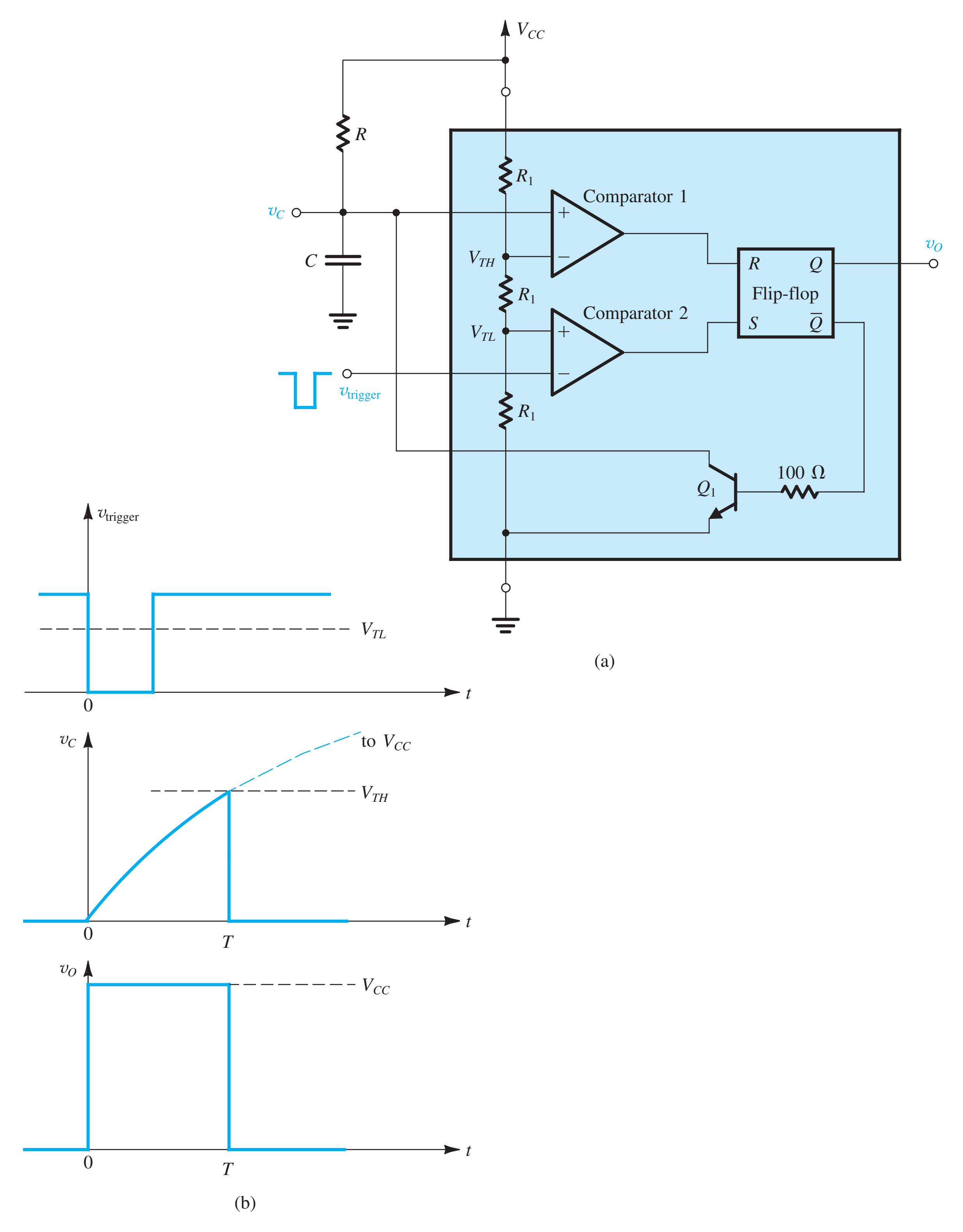


Figure 18.30 **(a)** The 555 timer connected to implement a monostable multivibrator. **(b)** Waveforms of the circuit in **(a)**.

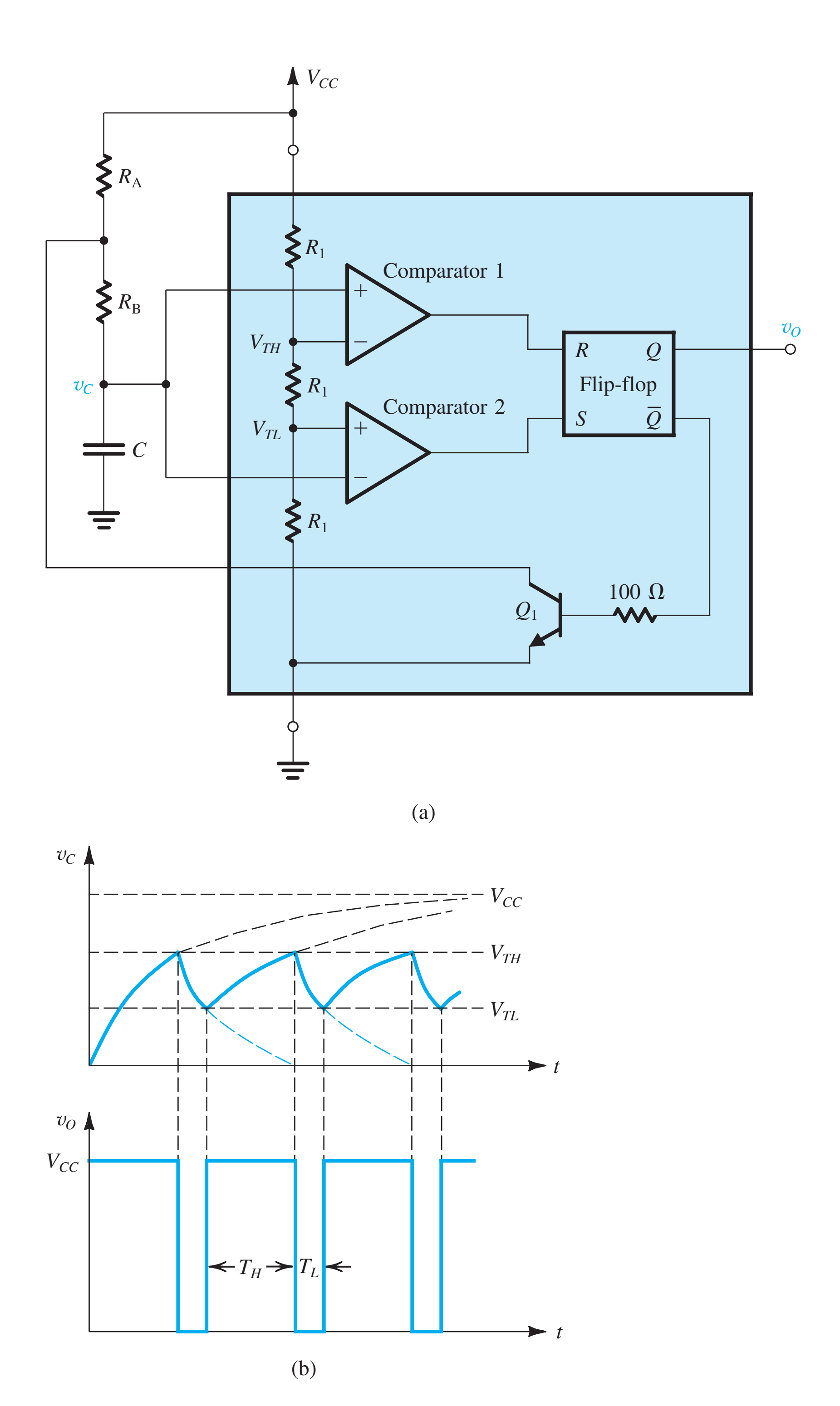


Figure 18.31 **(a)** The 555 timer connected to implement an astable multivibrator. **(b)** Waveforms of the circuit in **(a)**.

waveforms depicted in Fig. 18.31(b). Assume that initially C is discharged and the flip-flop is set. Thus v_O is high and Q_1 is off. Capacitor C will charge up through the series combination of R_A and R_B, and the voltage across it, v_C, will rise exponentially toward V_{CC}. As v_C crosses the level equal to V_{TL}, the output of comparator 2 goes low. This, however, has no effect on the circuit operation, and the flip-flop remains set. Indeed, this state continues until v_C reaches and begins to exceed the threshold of comparator 1, V_{TH}. At this instant of time, the output of comparator 1 goes high and resets the flip-flop. Thus v_O goes low, $\overline{Q}$ goes high, and transistor Q_1 is turned on. The saturated transistor Q_1 causes a voltage of approximately zero volts to appear at the common node of R_A and R_B. Thus C begins to discharge through R_B and the collector of Q_1. The voltage v_C decreases exponentially with a time constant CR_B toward 0 V. When v_C reaches the threshold of comparator 2, V_{TL}, the output of comparator 2, goes high and sets the flip-flop. The output v_O then goes high, and $\overline{Q}$ goes low, turning off Q_1. Capacitor C begins to charge through the series equivalent of R_A and R_B, and its voltage rises exponentially toward V_{CC} with a time constant $C(R_A + R_B)$. This rise continues until v_C reaches V_{TH}, at which time the output of comparator 1 goes high, resetting the flip-flop, and the cycle continues.

From the description above we see that the circuit of Fig. 18.31(a) oscillates and produces a square waveform at the output. The frequency of oscillation can be determined as follows. Reference to Fig. 18.31(b) indicates that the output will be high during the interval T_H, in which v_C rises from V_{TL} to V_{TH}. The exponential rise of v_C can be described by

$$v_C = V_{CC} - (V_{CC} - V_{TL})e^{-t/C(R_A+R_B)} \tag{18.42}$$

where $t = 0$ is the instant at which the interval T_H begins. Substituting $v_C = V_{TH} = \frac{2}{3}V_{CC}$ at $t = T_H$ and $V_{TL} = \frac{1}{3}V_{CC}$ results in

$$T_H = C(R_A + R_B)\ln 2 \simeq 0.69\, C(R_A + R_B) \tag{18.43}$$

We also note from Fig. 18.31(b) that v_O will be low during the interval T_L, in which v_C falls toward zero, from V_{TH} to V_{TL}. The exponential fall of v_C can be described by

$$v_C = V_{TH}e^{-t/CR_B} \tag{18.44}$$

where we have taken $t = 0$ as the beginning of the interval T_L. Substituting $v_C = V_{TL} = \frac{1}{3}V_{CC}$ at $t = T_L$ and $V_{TH} = \frac{2}{3}V_{CC}$ results in

$$T_L = CR_B \ln 2 \simeq 0.69\, CR_B \tag{18.45}$$

Equations (18.43) and (18.45) can be combined to obtain the period T of the output square wave as

$$T = T_H + T_L = 0.69\, C(R_A + 2R_B) \tag{18.46}$$

Also, the **duty cycle** of the output square wave can be found from Eqs. (18.43) and (18.45):

$$\text{Duty cycle} \equiv \frac{T_H}{T_H + T_L} = \frac{R_A + R_B}{R_A + 2R_B} \tag{18.47}$$

Note that the duty cycle will always be greater than 0.5 (50%); it approaches 0.5 if R_A is selected to be much smaller than R_B (unfortunately, at the expense of increased supply current).

EXERCISES

D18.23 Using a 10-nF capacitor C, find the value of R that yields an output pulse of 100 μs in the monostable circuit of Fig. 18.30(a).
Ans. 9.1 kΩ

D18.24 For the circuit in Fig. 18.31(a), with a 1-nF capacitor, find the values of R_A and R_B that result in an oscillation frequency of 100 kHz and a duty cycle of 75%.
Ans. 7.2 kΩ, 3.6 kΩ

18.8 Nonlinear Waveform-Shaping Circuits

Diodes or transistors can be combined with resistors to synthesize two-port networks having arbitrary nonlinear transfer characteristics. Such two-port networks can be employed in **waveform shaping**—that is, changing the waveform of an input signal in a prescribed manner to produce a waveform of a desired shape at the output. In this section we illustrate this application by a concrete example: the **sine-wave shaper**. This is a circuit whose purpose is to change the waveform of an input triangular-wave signal to a sine wave. Though simple, the sine-wave shaper is a practical building block used extensively in function generators. This method of generating sine waves should be contrasted to that using linear oscillators (Sections 18.1–18.3). Although linear oscillators produce sine waves of high purity, they are not convenient at very low frequencies. Also, linear oscillators are in general more difficult to tune over wide frequency ranges. In the following we discuss two distinctly different techniques for designing sine-wave shapers.

18.8.1 The Breakpoint Method

In the breakpoint method the desired nonlinear transfer characteristic (in our case the sine function shown in Fig. 18.32) is implemented as a piecewise linear curve. Diodes are utilized as switches that turn on at the various breakpoints of the transfer characteristic, thus switching into the circuit additional resistors that cause the transfer characteristic to change slope.

Consider the circuit shown in Fig. 18.33(a). It consists of a chain of resistors connected across the entire symmetrical voltage supply $+V$, $-V$. The purpose of this voltage divider is to generate reference voltages that will serve to determine the breakpoints in the transfer characteristic. In our example these reference voltages are denoted $+V_2$, $+V_1$, $-V_1$, $-V_2$. Note that the entire circuit is symmetrical, driven by a symmetrical triangular wave and generating a symmetrical sine-wave output. The circuit approximates each quarter-cycle of the sine wave by three straight-line segments; the breakpoints between these segments are determined by the reference voltages V_1 and V_2.

The circuit works as follows: Let the input be the triangular wave shown in Fig. 18.33(b), and consider first the quarter-cycle defined by the two points labeled 0 and 1. When the input signal is less in magnitude than V_1, none of the diodes conducts. Thus zero current flows through R_4, and the output voltage at B will be equal to the input voltage. But as the input

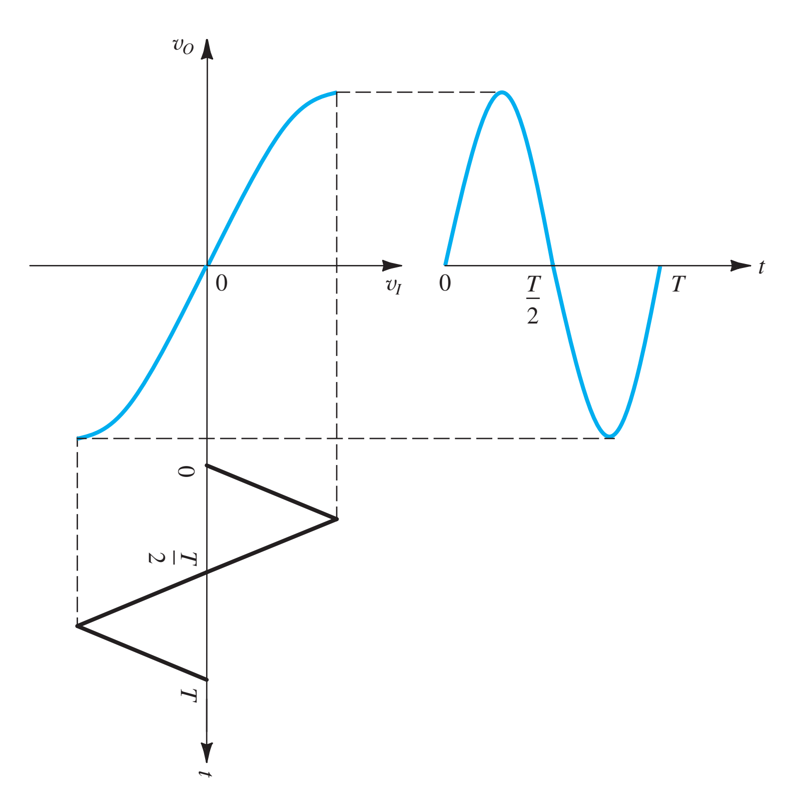


Figure 18.32 Using a nonlinear (sinusoidal) transfer characteristic to shape a triangular waveform into a sinusoid.

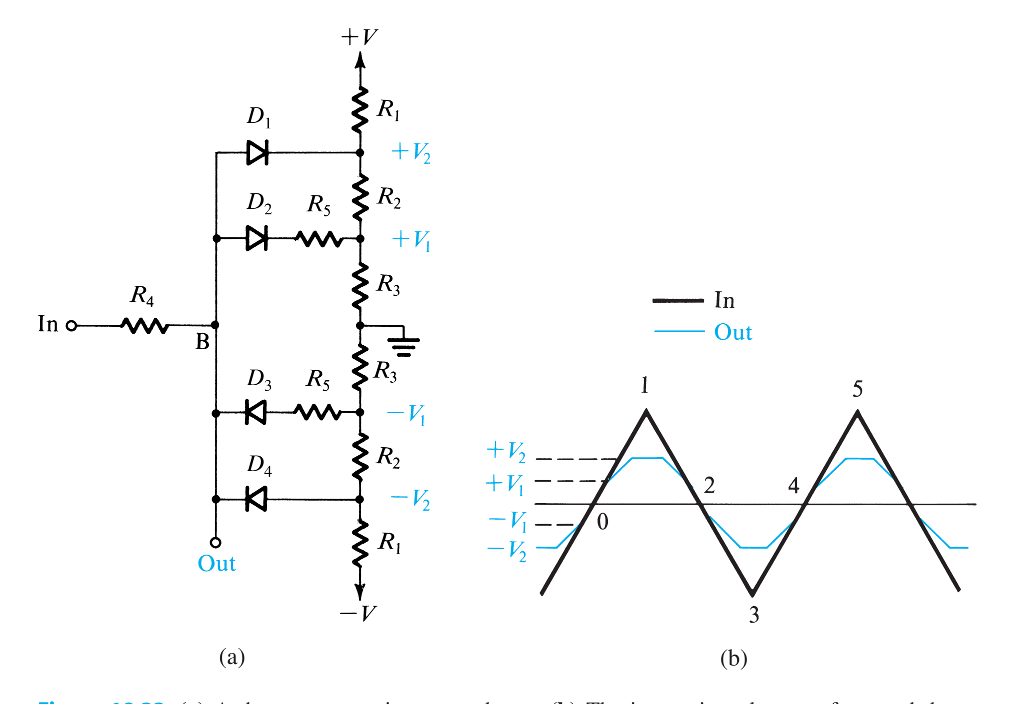


Figure 18.33 **(a)** A three-segment sine-wave shaper. **(b)** The input triangular waveform and the output approximately sinusoidal waveform.

rises to V_1 and above, D_2 (assumed ideal) begins to conduct. Assuming that the conducting D_2 behaves as a short circuit, we see that, for $v_I > V_1$,

$$v_O = V_1 + (v_I - V_1)\frac{R_5}{R_4 + R_5}$$

This implies that as the input continues to rise above V_1, the output follows, but with a reduced slope. This gives rise to the second segment in the output waveform, as shown in Fig. 18.33(b). Note that in developing the equation above we have assumed that the resistances in the voltage divider are low enough in value to cause the voltages V_1 and V_2 to be constant independent of the current coming from the input.

Next consider what happens as the voltage at point B reaches the second breakpoint determined by V_2. At this point, D_1 conducts, thus limiting the output v_O to V_2 (plus, of course, the voltage drop across D_1 if it is not assumed to be ideal). This gives rise to the third segment, which is flat, in the output waveform. The overall result is to "bend" the waveform and shape it into an approximation of the first quarter-cycle of a sine wave. Then, beyond the peak of the input triangular wave, as the input voltage decreases, the process unfolds, the output becoming progressively more like the input. Finally, when the input goes sufficiently negative, the process begins to repeat at $-V_1$ and $-V_2$ for the negative half-cycle.

Although the circuit is relatively simple, its performance is surprisingly good. A measure of goodness usually taken is to quantify the purity of the output sine wave by specifying the percentage **total harmonic distortion** (THD). This is the percentage ratio of the rms voltage of all harmonic components above the fundamental frequency (which is the frequency of the triangular wave) to the rms voltage of the fundamental (see also Chapter 12). Interestingly, one reason for the good performance of the diode shaper is the beneficial effects produced by the nonideal i–v characteristics of the diodes—that is, the exponential knee of the junction diode as it goes into forward conduction. The consequence is a relatively smooth transition from one line segment to the next.

Practical implementations of the breakpoint sine-wave shaper employ six to eight segments (compared with the three used in the example above). Also, transistors are usually employed to provide more versatility in the design, with the goal being increased precision and lower THD (see Grebene, 1984, pages 592–595).

18.8.2 The Nonlinear-Amplification Method

The other method we discuss for the conversion of a triangular wave into a sine wave is based on feeding the triangular wave to the input of an amplifier having a nonlinear transfer characteristic that approximates the sine function. One such amplifier circuit consists of a differential pair with a resistance connected between the two emitters, as shown in Fig. 18.34. With appropriate choice of the values of the bias current I and the resistance R, the differential amplifier can be made to have a transfer characteristic that closely approximates that shown in Fig. 18.32. Observe that for small v_I the transfer characteristic of the circuit of Fig. 18.34 is almost linear, as a sine waveform is near its zero crossings. At large values of v_I the nonlinear characteristics of the BJTs reduce the gain of the amplifier and cause the transfer characteristic to bend, approximating the sine wave as it approaches its peak. (More details on this circuit can be found in Grebene, 1984, pages 595–597.)

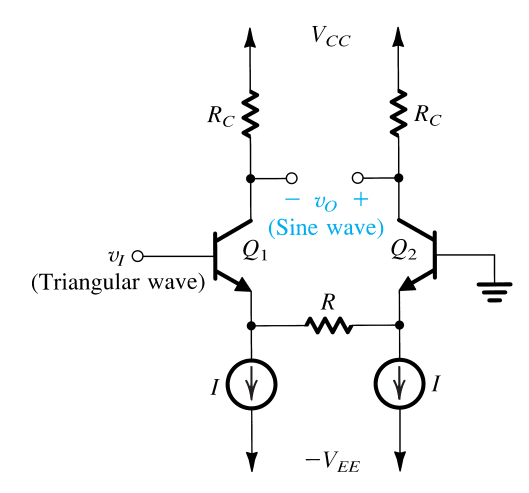


Figure 18.34 A differential pair with an emitter-degeneration resistance used to implement a triangular-wave to sine-wave converter. Operation of the circuit can be graphically described by Fig. 18.32.

EXERCISES

D18.25 The circuit in Fig. E18.25 is required to provide a three-segment approximation to the nonlinear i–v characteristic, $i = 0.1v^2$, where v is the voltage in volts and i is the current in milliamperes. Find the values of R_1, R_2, and R_3 such that the approximation is perfect at $v = 2$ V, 4 V, and 8 V. Calculate the error in current value at $v = 3$ V, 5 V, 7 V, and 10 V. Assume ideal diodes.

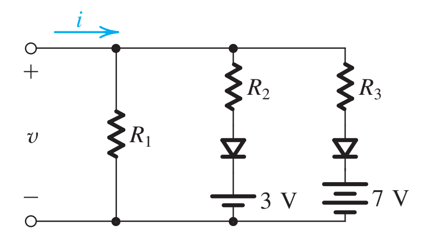


Figure E18.25

Ans. 5 kΩ, 1.25 kΩ, 1.25 kΩ; −0.3 mA, +0.1 mA, −0.3 mA, 0

18.26 A detailed analysis of the circuit in Fig. 18.34 shows that its optimum performance occurs when the values of I and R are selected so that $RI = 2.5V_T$, where V_T is the thermal voltage. For this design, the peak amplitude of the input triangular wave should be $6.6V_T$, and the corresponding sine wave across R has a peak value of $2.42V_T$. For $I = 0.25$ mA and $R_C = 10$ kΩ, find the peak amplitude of the sine-wave output v_O. Assume $\alpha \simeq 1$.
Ans. 4.84 V

Summary

- There are two distinctly different types of signal generator: the linear oscillator, which utilizes some form of resonance, and the nonlinear oscillator or function generator, which employs a switching mechanism implemented with a multivibrator circuit.
- A linear oscillator can be realized by placing a frequency-selective network in the feedback path of an amplifier (an op amp or a transistor). The circuit will oscillate at the frequency at which the total phase shift around the loop is zero or 360°, provided the magnitude of loop gain at this frequency is equal to, or greater than, unity.
- If in an oscillator the magnitude of loop gain is greater than unity, the amplitude will increase until a nonlinear amplitude-control mechanism is activated.
- The Wien-bridge oscillator, the phase-shift oscillator, the quadrature oscillator, and the active-filter-tuned oscillator are popular configurations for frequencies up to about 1 MHz. These circuits employ RC networks together with op amps or transistors. For higher frequencies, LC-tuned or crystal-tuned oscillators are utilized. Popular configurations include the Colpitts circuit for discrete-circuit implementation and the cross-coupled circuit for IC implementation at frequencies as high as hundreds of gigahertz.
- Crystal oscillators provide the highest possible frequency accuracy and stability.
- There are three types of multivibrator: bistable, monostable, and astable. Op-amp circuit implementations of multivibrators are useful in analog-circuit applications that require high precision.
- The bistable multivibrator has two stable states and can remain in either state indefinitely. It changes state when triggered. A comparator with hysteresis is bistable.
- A monostable multivibrator, also known as a one-shot, has one stable state, in which it can remain indefinitely. When triggered, it goes into a quasi-stable state in which it remains for a predetermined interval, thus generating, at its output, a pulse of known width.
- An astable multivibrator has no stable state. It oscillates between two quasi-stable states, remaining in each for a predetermined interval. It thus generates a periodic waveform at the output.
- A feedback loop consisting of an integrator and a bistable multivibrator can be used to generate triangular and square waveforms.
- The 555 timer, a commercially available IC, can be used with external resistors and a capacitor to implement high-quality monostable and astable multivibrators.
- A sine waveform can be generated by feeding a triangular waveform to a sine-wave shaper. A sine-wave shaper can be implemented either by using diodes (or transistors) and resistors, or by using an amplifier having a nonlinear transfer characteristic that approximates the sine function.

PROBLEMS

Section 18.1: Basic Principles of Sinusoidal Oscillators

18.1 Consider a sinusoidal oscillator consisting of an amplifier having a frequency-independent gain A (where A is positive) and a second-order bandpass filter with a pole frequency ω_0, a pole Q denoted Q, and a positive center-frequency gain K. Find the frequency of oscillation, and the condition that A and K must satisfy for sustained oscillation.

18.2 For the oscillator circuit described in Problem 18.1:

(a) Derive an expression for $d\phi/d\omega$, evaluated at $\omega = \omega_0$.

(b) Use the result of (a) to find an expression for the per-unit change in frequency of oscillation resulting from a phase-angle change of $\Delta\phi$, in the amplifier transfer function.

$$Hint: \frac{d}{dx}\left(\tan^{-1} y\right) = \frac{1}{1+y^2}\frac{dy}{dx}$$

18.3 For the oscillator described in Problem 18.1, show that, independent of the value of A and K, the poles of the circuit

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

lie at a radial distance of ω_0. Find the value of AK that results in poles appearing (a) on the $j\omega$ axis, and (b) in the right half of the s plane, at a horizontal distance from the $j\omega$ axis of $\omega_0/(2Q)$.

D 18.4 For the oscillator circuit in Fig. 18.3(a) find the percentage change in the oscillation frequency resulting from a change of +1% in the value of (a) L, (b) C, and (c) R.

18.5 An oscillator is formed by loading a transconductance amplifier having a positive gain with a parallel RLC circuit and connecting the output directly to the input (thus applying positive feedback with a factor $\beta = 1$). Let the transconductance amplifier have an input resistance of 5 kΩ and an output resistance of 5 kΩ. The LC resonator has $L = 1\ \mu$H, $C = 100$ pF, and $Q = 50$. For what value of transconductance G_m will the circuit oscillate? At what frequency?

18.6 In a particular oscillator characterized by the structure of Fig. 18.1, the frequency-selective network exhibits a loss of 12 dB and a phase shift of 180° at ω_0. Give the phase shift and the minimum gain that the amplifier must have for oscillation to begin.

18.7 An oscillator is designed by connecting in a loop three identical common-source amplifier stages of the type shown in Fig. P18.7. Note that the bias circuits are not shown, and assume that R and C include the transistor output resistance and capacitance, respectively. For the circuit to oscillate at a frequency ω_0, what must the phase angle provided by each amplifier stage be? Give an expression for ω_0. For sustained oscillations, what is the minimum g_m required of each transistor?

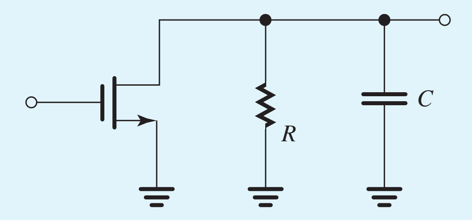


Figure P18.7

D 18.8 Consider the circuit of Fig. 18.4(a) with R_f removed to realize the comparator function. Find suitable values for all resistors so that the comparator output levels are ±3 V and the slope of the limiting characteristic is 0.05. Use power-supply voltages of ±5 V and assume the voltage drop of a conducting diode to be 0.7 V.

D 18.9 Consider the circuit of Fig. 18.4(a) with R_f removed to realize the comparator function. Sketch the transfer characteristic. Show that by connecting a dc source V_B to the virtual ground of the op amp through a resistor R_B, the transfer characteristic is shifted along the v_I axis to the point $v_I = -(R_1/R_B)V_B$. Utilizing available ±5-V dc supplies for $\pm V$ and for V_B, find suitable component values so that the limiting levels are ±3 V and the comparator threshold is at $v_I = +2$ V. Neglect the diode voltage drop (i.e., assume that $V_D = 0$). The input resistance of the comparator is to be 100 kΩ, and the slope in the limiting regions is to be ≤ 0.05 V/V. Use standard 5% resistors (see Appendix J).

18.10 Denoting the zener voltages of Z_1 and Z_2 by V_{Z1} and V_{Z2} and assuming that in the forward direction the voltage drop

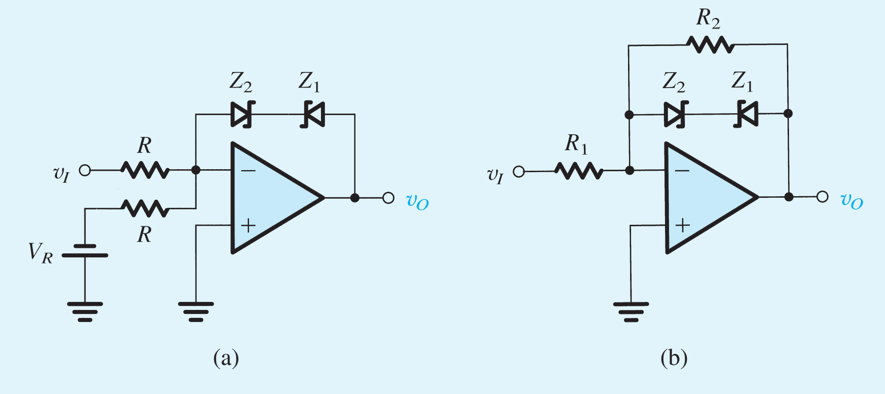


Figure P18.10

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

is approximately 0.7 V, sketch and clearly label the transfer characteristics v_O–v_I of the circuits in Fig. P18.10. Assume the op amps to be ideal.

Section 18.2: Op Amp–RC Oscillator Circuits

18.11 For the Wien-bridge oscillator circuit in Fig. 18.5, show that the transfer function of the feedback network $[V_a(s)/V_o(s)]$ is that of a bandpass filter. Find ω_0 and Q of the poles, and find the center-frequency gain.

18.12 For the Wien-bridge oscillator of Fig. 18.5, let the closed-loop amplifier (formed by the op amp and the resistors R_1 and R_2) exhibit a phase shift of $-3°$ in the neighborhood of $\omega = 1/CR$. Find the frequency at which oscillations can occur in this case in terms of CR. (*Hint:* Use Eq. 18.11.)

18.13 For the Wien-bridge oscillator of Fig. 18.5, use the expression for loop gain in Eq. (18.10) to find the poles of the closed-loop system. Give the expression for the pole Q, and use it to show that to locate the poles in the right half of the s plane, R_2/R_1 must be selected to be greater than 2.

D 18.14 Reconsider Exercise 18.5 with R_3 and R_6 increased to reduce the output voltage. What values are required for a peak-to-peak output of 8 V? What results if R_3 and R_6 are open-circuited?

18.15 For the circuit in Fig. P18.15, find $L(s)$, $L(j\omega)$, the frequency for zero loop phase, and R_2/R_1 for oscillation. Assume the op amp to be ideal.

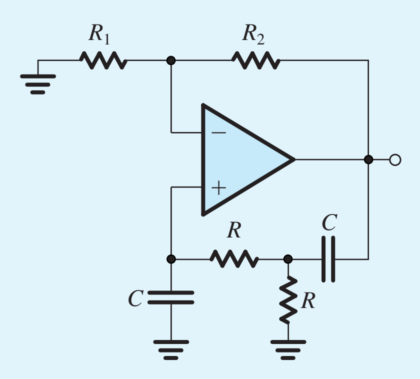


Figure P18.15

18.16 Repeat Problem 18.15 for the circuit in Fig. P18.16.

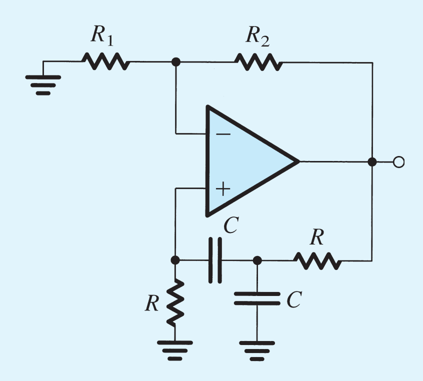


Figure P18.16

***18.17** Consider the circuit of Fig. 18.7 with the 50-kΩ potentiometer replaced by two fixed resistors: 10 kΩ between the op amp's negative input and ground, and 15 kΩ. Modeling each diode as a 0.65-V battery in series with a 100-Ω resistance, find the peak-to-peak amplitude of the output sinusoid.

D **18.18 Design the circuit of Fig. 18.7 for operation at 10 kHz using $R = 10$ kΩ. If at 10 kHz the op amp provides an excess phase shift (lag) of 5.7°, what will be the frequency of oscillation? (Assume that the phase shift introduced by the op amp remains constant for frequencies around 10 kHz.) To restore operation to 10 kHz, what change must be made in the shunt resistor of the Wien bridge? Also, to what value must R_2/R_1 be changed?

***18.19** For the circuit of Fig. 18.9, connect an additional resistor ($R = 10$ kΩ) in series with the rightmost capacitor C. For this modification (and ignoring the amplitude stabilization circuitry), find the loop gain $A\beta$ by breaking the circuit at node X. Find R_f for oscillation to begin, and find f_0.

D 18.20 For the circuit in Fig. P18.19, break the loop at node X and find the loop gain (working backward for simplicity to

SIM = Multisim/PSpice; * = difficult problem; ** = more difficult; *** = very challenging; D = design problem

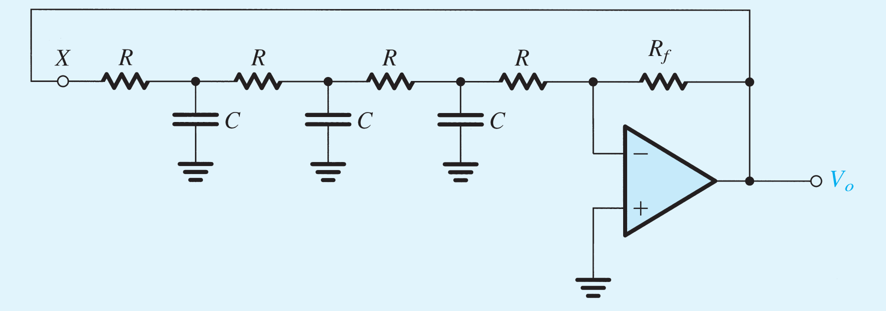


Figure P18.20

find V_x in terms of V_o). For $R = 10\ \text{k}\Omega$, find C and R_f to obtain sinusoidal oscillations at 15 kHz.

***18.21** Consider the quadrature-oscillator circuit of Fig. 18.10 without the limiter. Let the resistance R_f be equal to $2R/(1+\Delta)$, where $\Delta \ll 1$. Show that the poles of the characteristic equation are in the right-half s plane and given by $s \simeq (1/CR)[(\Delta/4) \pm j]$.

D 18.22 Using $C = 1.6$ nF, find the value of R such that the circuit of Fig. 18.12 produces 10-kHz sine waves. If the diode drop is 0.7 V, find the peak-to-peak amplitude of the output sine wave. How do you modify the circuit to double the output amplitude? (*Hint*: A square wave with peak-to-peak amplitude of V volts has a fundamental component with $4V/\pi$ volts peak-to-peak amplitude.)

***18.23** Assuming that the diode-clipped waveform in Exercise 18.9 is nearly an ideal square wave and that the resonator Q is 20, provide an estimate of the distortion in the output sine wave by calculating the magnitude (relative to the fundamental) of

(a) the second harmonic
(b) the third harmonic
(c) the fifth harmonic
(d) the rms of harmonics to the tenth

Note that a square wave of amplitude V and frequency ω is represented by the series

$$\frac{4V}{\pi}\left(\sin\omega t + \frac{1}{3}\sin 3\omega t + \frac{1}{5}\sin 5\omega t + \frac{1}{7}\sin 7\omega t + \ldots\right)$$

Section 18.3: LC and Crystal Oscillators

18.24 For the Colpitts oscillator circuit in Fig. P18.24, derive an equation governing circuit operation and hence find the frequency of oscillation and the condition on the gain $g_m R_L$ that ensures that oscillations will start. Assume that R_L includes r_o of Q_1. Simplify your final expressions by assuming r_π is large. Observe that this circuit is based on the configuration in Fig. 18.13(a) except that here the biasing circuit is included and the collector is placed at signal ground.

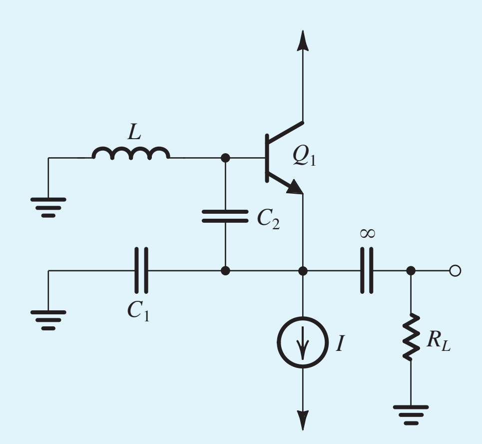


Figure P18.24

18.25 For the Colpitts oscillator circuit in Fig. P18.25, derive an equation governing circuit operation and hence find the frequency of oscillation and the condition the gain $g_m R_L$ must

satisfy for oscillations to start. Assume that R_L includes the MOSFET's r_o.

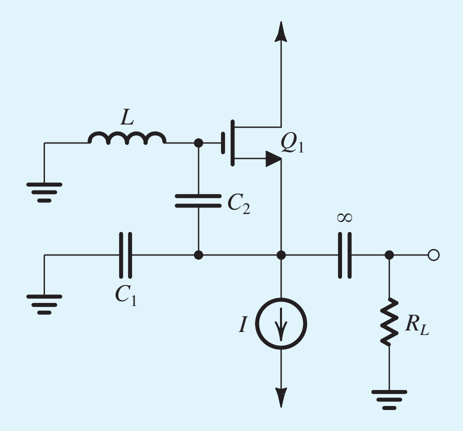


Figure P18.25

18.26 For the Colpitts oscillator circuit in Fig. P18.26, derive an equation governing circuit operation and hence find the frequency of oscillation and the condition the gain $g_m R_L$ must satisfy to ensure that oscillations will start. Neglect r_o of the BJT. Simplify your final expressions by assuming that r_π is large. Note that this circuit is based on the configuration of Fig. 18.13(a) but with the bias circuit included and the base grounded.

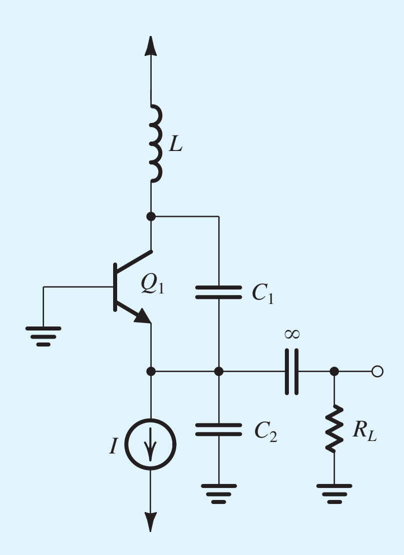


Figure P18.26

18.27 For the Colpitts oscillator circuit in Fig. P18.27, derive an equation governing circuit operation and hence find the frequency of oscillation and the condition the gain $g_m R_L$ must satisfy to ensure that oscillations will start. Assume that r_o of the BJT is included in R_L and neglect R_f (i.e., assume $R_f \gg \omega_0 L$). Simplify your final expressions by assuming r_π is large. Observe that this circuit is similar to that in Fig. 18.15 except for utilizing a different biasing scheme.

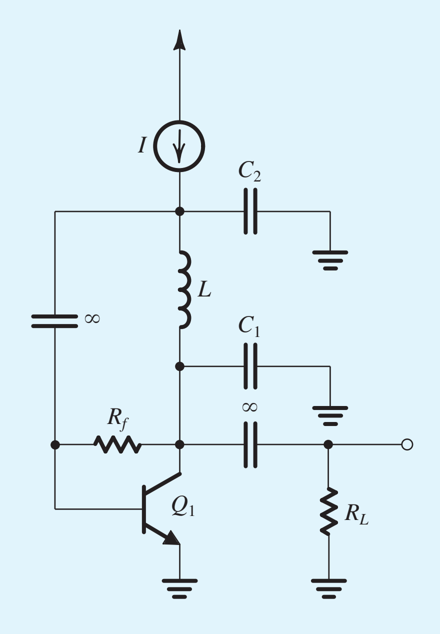


Figure P18.27

***18.28** The LC oscillator in Fig. P18.28 is based on connecting a positive-gain amplifier (formed by Q_1, Q_2, and R_C) with a bandpass RLC circuit in a feedback loop.

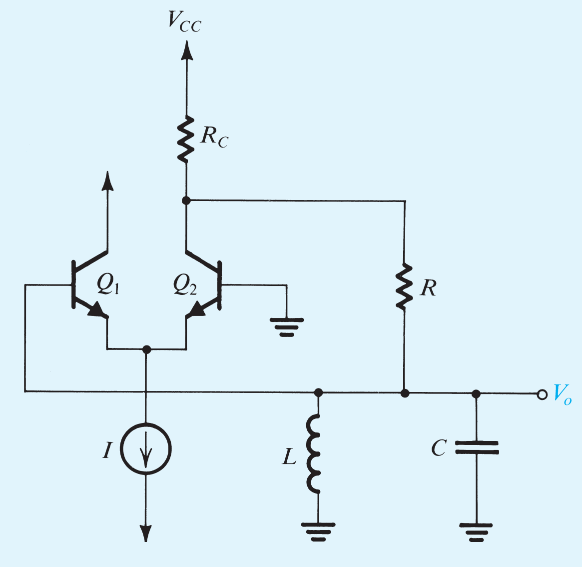


Figure P18.28

(a) Replace the BJTs with their small-signal models while neglecting r_π and r_o (to simplify matters).
(b) By inspection of the circuit found in (a), find the frequency of oscillation and the condition required for oscillations to start. Express the latter as the minimum required value of (IR_C).
(c) If IR_C is selected equal to 1 V, show that oscillations will start. If oscillations grow to the point that V_o is large enough to turn the BJTs on and off, show that the signal at the collector of Q_2 will be a square wave of 1 V peak-to-peak. Estimate the peak-to-peak amplitude of the output sine wave V_o.

D 18.29 Design the cross-coupled LC oscillator of Fig. 18.16(a) to operate at $\omega_0 = 20$ Grad/s. The IC inductors available have $L = 5$ nH and $Q = 10$. If the transistor $r_o = 5$ kΩ, find the required value of C and the minimum required value of g_m at which Q_1 and Q_2 are to be operated.

18.30 Consider the Pierce crystal oscillator of Fig. 18.18 with the crystal as specified in Exercise 18.13. Let C_1 be variable in the range 1 pF to 10 pF, and let C_2 be fixed at 10 pF. Find the range over which the oscillation frequency can be tuned. (*Hint:* Use the result in the statement leading to the expression in Eq. 18.29.)

Section 18.4: Bistable Multivibrators

D 18.31 Design the bistable circuit in Fig. 18.21(a) to obtain a hysteresis of 2-V width. The op amp saturates at ±5 V. Select $R_1 = 10$ kΩ and determine R_2.

18.32 Consider the bistable circuit of Fig. 18.21(a) with the op amp's positive input terminal connected to a positive-voltage source V through a resistor R_3.

(a) Derive expressions for the threshold voltages V_{TL} and V_{TH} in terms of the op amp's saturation levels L_+ and L_-, R_1, R_2, R_3, and V.
(b) Let $L_+ = -L_- = 10$ V, $V = 15$ V, and $R_1 = 10$ kΩ. Find the values of R_2 and R_3 that result in $V_{TL} = +4.9$ V and $V_{TH} = +5.1$ V.

18.33 Consider the bistable circuit of Fig. 18.22(a) with the op amp's negative-input terminal disconnected from ground and connected to a reference voltage V_R.

(a) Derive expressions for the threshold voltages V_{TL} and V_{TH} in terms of the op amp's saturation levels L_+ and L_-, R_1, R_2, and V_R.
(b) Let $L_+ = -L_- = V$ and $R_1 = 10$ kΩ. Find R_2 and V_R that result in threshold voltages of 0 and $V/10$.

18.34 For the circuit in Fig. P18.34, sketch and label the transfer characteristic v_O–v_I. The diodes are assumed to have a constant 0.7-V drop when conducting, and the op amp saturates at ±12 V. What is the maximum diode current?

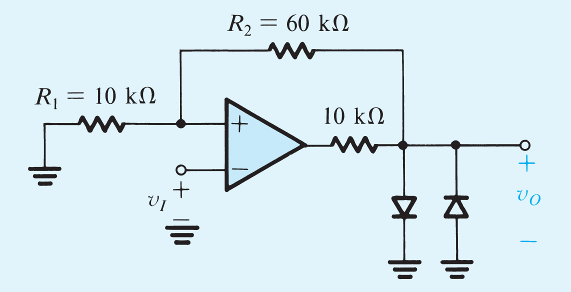


Figure P18.34

18.35 Consider the circuit of Fig. P18.34 with R_1 eliminated and R_2 short-circuited. Sketch and label the transfer characteristic v_O–v_I. Assume that the diodes have a constant 0.7-V drop when conducting and that the op amp saturates at ±12 V.

***18.36** Consider a bistable circuit having a noninverting transfer characteristic with $L_+ = -L_- = 12$ V, $V_{TL} = -1$ V, and $V_{TH} = +1$ V.

(a) For a 0.5-V-amplitude sine-wave input having zero average, what is the output?
(b) Describe the output if a sinusoid of frequency f and amplitude of 1.1 V is applied at the input. By how much can the average of this sinusoidal input shift before the output becomes a constant value?

D 18.37 Design the circuit of Fig. 18.25(a) to realize a transfer characteristic with ±7.5-V output levels and ±7.5-V threshold values. Design so that when $v_I = 0$ V a current of 0.5 mA flows in the feedback resistor and a current of 1 mA flows through the zener diodes. Assume that the output saturation levels of the op amp are ±10 V. Specify the voltages of the zener diodes and give the values of all resistors.

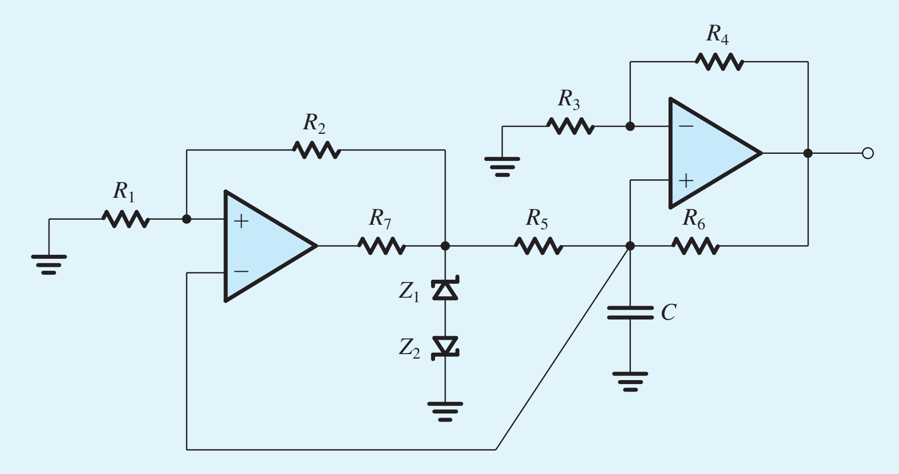


Figure P18.41

Section 18.5: Generation of Square and Triangular Waveforms Using Astable Multivibrators

18.38 Find the frequency of oscillation of the circuit in Fig. 18.26(b) for the case $R_1 = 10$ kΩ, $R_2 = 16$ kΩ, $C = 5$ nF, and $R = 62$ kΩ.

D 18.39 Augment the astable multivibrator circuit of Fig. 18.26(b) with an output limiter of the type shown in Fig. 18.25(b). Design the circuit to obtain an output square wave with 5-V amplitude and 1-kHz frequency using a 10-nF capacitor C. Use $\beta = 0.462$, and design for a current in the resistive divider approximately equal to the average current in the RC network over a half-cycle. Assuming ±13-V op-amp saturation voltages, arrange for the zener to operate at a minimum current of 1 mA. Specify the values of all resistors and the zenor voltage.

D 18.40 Using the scheme of Fig. 18.27, design a circuit that provides square waves of 10 V peak to peak and triangular waves of 10 V peak to peak. The frequency is to be 1 kHz. Implement the bistable circuit with the circuit of Fig. 18.25(b). Use a 0.01-μF capacitor and specify the values of all resistors and the required zener voltage. Design for a minimum zener current of 1 mA and for a maximum current in the resistive divider of 0.2 mA. Assume that the output saturation levels of the op amps are ±12 V.

D *18.41 The circuit of Fig. P18.41 consists of an inverting bistable multivibrator with an output limiter and a noninverting integrator. Using equal values for all resistors except R_7 and a 0.5-nF capacitor, design the circuit to obtain a square wave at the output of the bistable multivibrator of 15-V peak-to-peak amplitude and 10-kHz frequency. Sketch and label the waveform at the integrator output. Assuming ±13-V op-amp saturation levels, design for a minimum zener current of 1 mA. Specify the zener voltage required, and give the values of all resistors.

Section 18.6: Generation of a Standardized Pulse—The Monostable Multivibrator

D 18.42 For the monostable circuit considered in Exercise 18.22, calculate the recovery time.

***18.43** Figure P18.43 shows a monostable multivibrator circuit. In the stable state, $v_O = L_+$, $v_A = 0$, and $v_B = -V_{\text{ref}}$. The circuit can be triggered by applying a positive input pulse of height greater than V_{ref}. For normal operation, $C_1R_1 \ll CR$. Show the resulting waveforms of v_O and v_A. Also, show that the pulse generated at the output will have a width T given by

$$T = CR\ln\left(\frac{L_+ - L_-}{V_{\text{ref}}}\right)$$

Note that this circuit has the interesting property that the pulse width can be controlled by changing V_{ref}.

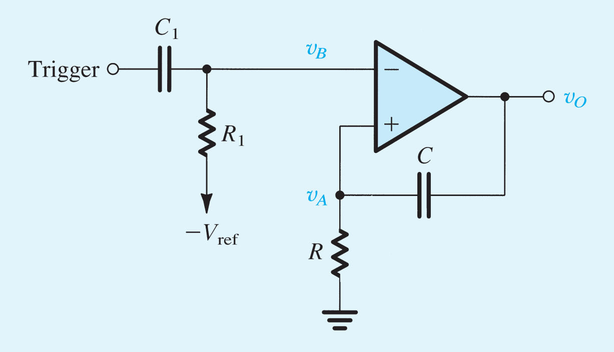


Figure P18.43

D *18.44 Using the circuit of Fig. 18.28, with a nearly ideal op amp for which the saturation levels are ±13 V, design a monostable multivibrator to provide a negative output pulse of 100-μs duration. Use capacitors of 0.1 nF and 1 nF. Wherever possible, choose resistors of 100 kΩ in your design. Diodes have a drop of 0.7 V. What is the minimum input step size that will ensure triggering? How long does the circuit take to recover to a state in which retriggering is possible with a normal output?

Section 18.7: Integrated-Circuit Timers

18.45 Consider the 555 circuit of Fig. 18.29 when the Threshold and the Trigger input terminals are joined together and connected to an input voltage v_I. Verify that the transfer characteristic v_O–v_I is that of an inverting bistable circuit with thresholds $V_{TL} = \frac{1}{3}V_{CC}$ and $V_{TH} = \frac{2}{3}V_{CC}$ and output levels of 0 and V_{CC}.

D 18.46 (a) Using a 0.5-nF capacitor C in the circuit of Fig. 18.30(a), find the value of R that results in an output pulse of 10-μs duration.
(b) If the 555 timer used in (a) is powered with $V_{CC} = 12$ V, and assuming that V_{TH} can be varied externally (i.e., it need not remain equal to $\frac{2}{3}V_{CC}$), find its required value so that the pulse width is increased to 20 μs, with other conditions the same as in (a).

D 18.47 Using a 680-pF capacitor, design the astable circuit of Fig. 18.31(a) to obtain a square wave with a 20-kHz frequency and an 80% duty cycle. Specify the values of R_A and R_B.

***18.48** The node in the 555 timer at which the voltage is V_{TH} (i.e., the inverting input terminal of comparator 1) is usually connected to an external terminal. This allows the user to change V_{TH} externally (i.e., V_{TH} no longer remains at $\frac{2}{3}V_{CC}$). Note, however, that whatever the value of V_{TH} becomes, V_{TL} always remains $\frac{1}{2}V_{TH}$.

(a) For the astable circuit of Fig. 18.31, rederive the expressions for T_H and T_L, expressing them in terms of V_{TH} and V_{TL}.
(b) For the case $C = 1$ nF, $R_A = 7.2$ kΩ, $R_B = 3.6$ kΩ, and $V_{CC} = 5$ V, find the frequency of oscillation and the duty cycle of the resulting square wave when no external voltage is applied to the terminal V_{TH}.
(c) For the design in (b), let a sine-wave signal of a much lower frequency than that found in (b) and of 1-V peak amplitude be capacitively coupled to the circuit node V_{TH}. This signal will cause V_{TH} to change around its quiescent value of $\frac{2}{3}V_{CC}$, and thus T_H will change correspondingly—a modulation process. Find T_H, and find the frequency of oscillation and the duty cycle at the two extreme values of V_{TH}.

Section 18.8: Nonlinear Waveform-Shaping Circuits

D *18.49 The two-diode circuit shown in Fig. P18.49 can provide a crude approximation to a sine-wave output when driven by a triangular waveform. To obtain a good approximation, we select the peak of the triangular waveform, V, so that the slope of the desired sine wave at the zero crossings is equal to that of the triangular wave. Also, the value of R is selected so that when v_I is at its peak, the output voltage is equal to the desired peak of the sine wave. If the diodes exhibit a voltage drop of 0.7 V at 1-mA current, changing at the rate of 0.1 V per decade, find the values of V and R that will yield an approximation to a sine waveform of 0.7-V peak amplitude. Then find the angles θ (where $\theta = 90°$ when v_I is at its peak) at which the output of the circuit, in volts, is 0.7, 0.65, 0.6, 0.55, 0.5, 0.4, 0.3, 0.2, 0.1, and 0. Use the

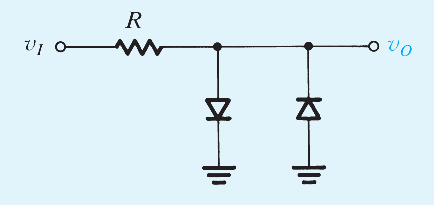


Figure P18.49

angle values obtained to determine the values of the exact sine wave (i.e., 0.7 sin θ), and thus find the percentage error of this circuit as a sine shaper. Provide your results in tabular form.

D 18.50 Design a two-segment sine-wave shaper using a 6.8-kΩ-input resistor, two diodes, and two clamping voltages. The circuit, fed by an 8-V peak-to-peak triangular wave, should limit the amplitude of the output signal via a 0.7-V diode to a value corresponding to that of a sine wave whose zero-crossing slope matches that of the triangle. What are the clamping voltages you have chosen?

18.51 Show that the output voltage of the circuit in Fig. P18.51 is given by

$$v_O = -V_T \ln\left(\frac{v_I}{I_S R}\right), \quad v_I > 0$$

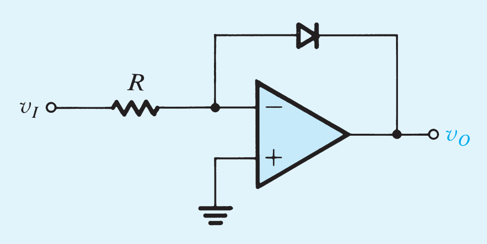


Figure P18.51

where I_S is the saturation current of the diode and V_T is the thermal voltage. Since the output voltage is proportional to the logarithm of the input voltage, the circuit is known as a **logarithmic amplifier**. Such amplifiers find application in situations where it is desired to compress the signal range.

18.52 Verify that the circuit in Fig. P18.52 implements the transfer characteristic $v_O = -v_1 v_2$ for $v_1, v_2 > 0$. Such a circuit is known as an analog multiplier. Check the circuit's performance for various combinations of input voltage of values, say, 0.5 V, 1 V, 2 V, and 3 V. Assume all diodes to be identical, with 700-mV drop at 1-mA current. Note that a *squarer* can easily be produced using a single input (e.g., v_1) connected via a 0.5-kΩ resistor (rather than the 1-kΩ resistor shown).

****18.53** Detailed analysis of the circuit in Fig. 18.34 shows that optimum performance (as a sine shaper) occurs when the values of I and R are selected so that $RI = 2.5V_T$, where V_T is the thermal voltage, and the peak amplitude of the input triangular wave is $6.6V_T$. If the output is taken across R (i.e., between the two emitters), find v_I corresponding to $v_O = 0.25V_T$, $0.5V_T$, V_T, $1.5V_T$, $2V_T$, $2.4V_T$, and $2.42V_T$. Plot v_O–v_I and compare to the ideal curve given by

$$v_O = 2.42V_T \sin\left(\frac{v_I}{6.6V_T} \times 90°\right)$$

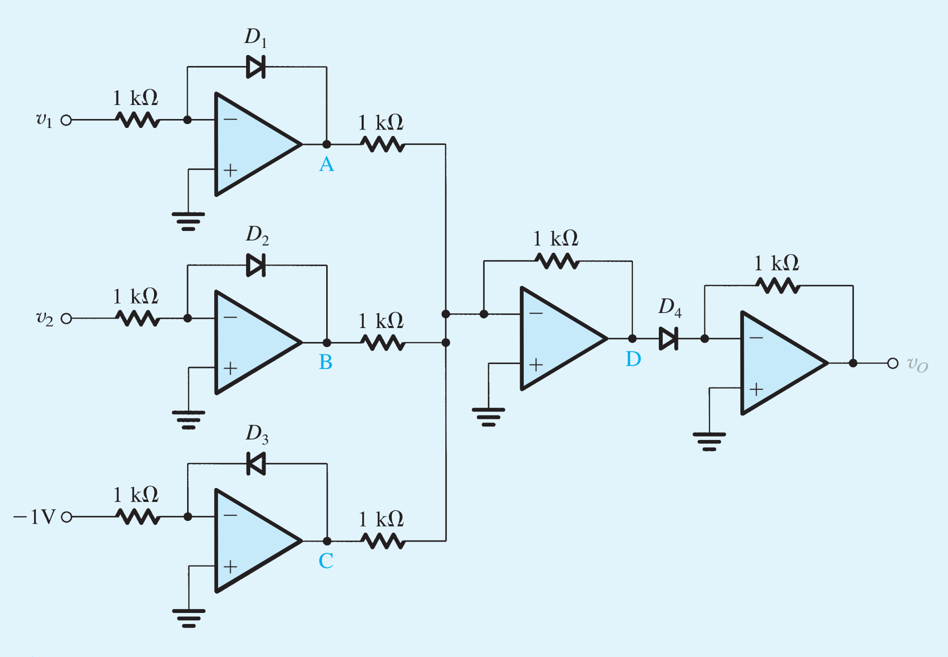


Figure P18.52

APPENDICES ON COMPANION WEBSITE

For your convenience, ten additional chapters on important reference topics are included on the Companion Website. In PDF format, the Appendices are fully searchable and can be bookmarked.

Appendix A: VLSI Fabrication Technology This article is a concise explanation of the technology that goes into fabricating integrated circuits. The different processes used are described and compared, and the characteristics of the resulting devices presented. Design considerations that restrict IC designers are explored.

Appendix B: SPICE Device Models and Design and Simulation Examples Using PSpice® and Multisim™ This three-part appendix could stand as a book on its own. Part 1 describes the models SPICE programs use to represent op amps, diodes, MOSFETs, and BJTs in integrated circuits. A thorough understanding of these models is critical for designers trying to extract meaningful information from an analysis. Part 2 describes and discusses all the PSpice® simulations, while Part 3 does the same for the Multisim™ simulations. This is a rich resource to help analyze, experiment with, and design circuits that relate to the topics studied in *Microelectronic Circuits.*

Appendix C: Two-Port Network Parameters Throughout the text, we use different possible ways to characterize linear two-port networks. This appendix summarizes the y, *z, h,* and *g* parameters and provides their equivalent-circuit representations.

Appendix D: Some Useful Network Theorems This article reviews Thévenin's theorem, Norton's theorem, and the source-absorption theorem, all of which are useful in simplifying the analysis of electronic circuits.

Appendix E: Single-Time-Constant Circuits STC circuits are composed of, or can be reduced to, one reactive component (inductance or capacitance) and one resistance. This is important to the design and analysis of linear and digital circuits. Analyzing an amplifier circuit can usually be reduced to the analysis of one or more STC circuits.

Appendix F: *s*-Domain Analysis: Poles, Zeroes, and Bode Plots Most of the work in analyzing the frequency response of an amplifier involves finding the amplifier voltage gain as a function of the complex frequency *s*. The tools to do this are summarized in this appendix.

Appendix G: Comparison of the MOSFET and the BJT Provides a comprehensive compilation and comparison of the properties of the MOSFET and the BJT. The comparison is aided by the inclusion of typical parameter values of devices fabricated with modern process technologies.

Appendix H: Design of Stagger-Tuned Amplifiers Provides a systematic procedure for the design of this type of tuned amplifier (Section 17.12.6).

Appendix I: Bibliography An excellent resource for students beginning research projects, this bibliography outlines key reference works on electronic circuits, circuit and system analysis, devices and IC fabrication, op amps, analog and digital circuits, filters and tuned amplifiers, and SPICE.

Appendix L: Answers to Selected Problems.

APPENDIX A

VLSI FABRICATION TECHNOLOGY

Introduction

Since the first edition of this text, we have witnessed a fantastic evolution in **VLSI** (very-large-scale integrated circuits) technology. In the late 1970s, non-self-aligned metal gate MOSFETs with gate lengths in the order of 10 μm were the norm. Current VLSI fabrication technology is already at the physical scaling limit with gate lengths in the 20-nm regime. This represents a reduction in device size of almost 1000x, along with an even more impressive increase in the number of devices per VLSI chip. Future development in VLSI technology must rely on new device concepts and new materials, taking quantum effects into account. While this is a very exciting time for researchers to explore new technology, we can also be assured that the "traditional" **CMOS** and **BiCMOS** (bipolar CMOS) fabrication technology will continue to be the workhorse of the microelectronic industry for many more years to come.

The purpose of this appendix is to familiarize the reader with VLSI fabrication technology. Brief explanations of standard VLSI processing steps are given. The variety of devices available in CMOS and BiCMOS fabrication technologies are also presented. In particular, the availability of components in the **IC** (integrated circuit) environment that are distinct from discrete circuit design will be discussed. In order to enjoy the economics of integrated circuits, designers have to overcome some serious device limitations (such as poor device tolerances) while exploiting device advantages (such as good component matching). An understanding of device characteristics is therefore essential in designing high-performance custom VLSIs.

This appendix will consider only silicon-based (Si) technologies. Although other compound materials in groups III through V, such as gallium arsenide (GaAs) and aluminum gallium nitride (AlGaN), are also used to implement VLSI chips, silicon is still the most popular material, with excellent cost–performance trade-off. Recent development in SiGe and strained-silicon technologies will further strengthen the position of Si-based fabrication processes in the microelectronic industry for many more years to come.

Silicon is an abundant element and occurs naturally in the form of sand. It can be refined using well-established purification and crystal growth techniques. It also exhibits suitable physical properties for fabricating active devices with good electrical characteristics. In addition, silicon can be easily oxidized to form an excellent insulator, SiO_2 (glass). This native oxide is useful for constructing capacitors and MOSFETs. It also serves as a diffusion barrier that can mask against unwanted impurities from diffusing into the high-purity silicon material. This masking property allows the electrical properties of the silicon to be altered in predefined areas. Therefore, active and passive elements can be built on the same piece of material (substrate). The components can then be interconnected using metal layers (similar to those used in printed-circuit boards) to form a monolithic IC.

A.1 IC Fabrication Steps

The basic IC fabrication steps will be described in the following sections. Some of these steps may be carried out many times, in different combinations and/or processing conditions during a complete fabrication run.

A.1.1 Silicon Wafers

The starting material for modern integrated circuits is very-high-purity, single-crystal silicon. The material is initially grown as a single crystal ingot. It takes the shape of a steel-gray solid cylinder 10 cm to 30 cm in diameter and can be one to two meters in length. This crystal is then sawed (like a loaf of bread) to produce circular **wafers** that are 400 μm to 600 μm thick (a micrometer, or micron, μm, is a millionth of a meter). The surface of the wafer is then polished to a mirror finish using chemical and mechanical polishing (CMP) techniques. Semiconductor manufacturers usually purchase ready-made silicon wafers from a supplier and rarely start their fabrication process in ingot form.

The basic electrical and mechanical properties of the wafer depend on the orientation of the crystalline structure, the impurity concentrations, and the type of impurities present. These variables are strictly controlled during crystal growth. A specific amount of impurities can be added to the pure silicon in a process known as doping. This allows the alteration of the electrical properties of the silicon, in particular its resistivity. Depending on the types of impurity, either holes (in ***p*-type** silicon) or electrons (in ***n*-type** silicon) can be responsible for electrical conduction. If a large number of impurity atoms is added, the silicon will be heavily doped (e.g., concentration $> \sim 10^{18}$ atoms/cm^{-3}). When designating the relative doping concentrations in semiconductor material, it is common to use the + and – symbols. A heavily doped (low-resistivity) n-type silicon wafer is referred to as $n+$ material, while a lightly doped material (e.g., concentration $< \sim 10^{16}$ atoms/cm^{-3}) is referred to as $n-$. Similarly, $p+$ and $p-$ designations refer to the heavily doped and lightly doped p-type regions, respectively. The ability to control the type of impurities and the doping concentration in the silicon permits the formation of diodes, transistors, and resistors in integrated circuits.

A.1.2 Oxidation

In **oxidation**, silicon reacts with oxygen to form silicon dioxide (SiO_2). To speed up this chemical reaction, it is necessary to carry out the oxidation at high temperatures (e.g., 1000–1200°C) and inside ultraclean furnaces. To avoid the introduction of even small quantities of contaminants (which could significantly alter the electrical properties of the silicon), it is necessary to operate in a **clean room** . Particle filters are used to ensure that the airflow in the processing area is free from dust. All personnel must protect the clean-room environment by wearing special lint-free clothing that covers a person from head to toe.

The oxygen used in the reaction can be introduced either as a high-purity gas (referred to as a "**dry oxidation**") or as steam (forming a "**wet oxidation**"). In general, wet oxidation has a faster growth rate, but dry oxidation gives better electrical characteristics. The thermally grown oxide layer has excellent electrical insulation properties. The dielectric strength for **SiO_2** is approximately 10^7 V/cm. It has a dielectric constant of about 3.9, and it can be used to form excellent MOS capacitors. Silicon dioxide can also serve as an effective mask against many impurities, allowing the introduction of dopants into the silicon only in regions that are not covered with oxide.

Silicon dioxide is a transparent film, and the silicon surface is highly reflective. If white light is shone on an oxidized wafer, constructive and destructive interference will cause

certain colors to be reflected. The wavelengths of the reflected light depend on the thickness of the oxide layer. In fact, by categorizing the color of the wafer surface, one can deduce the thickness of the oxide layer. The same principle is used by more sophisticated optical inferometers to measure film thickness. On a processed wafer, there will be regions with different oxide thicknesses. The colors can be quite vivid and are immediately obvious when a finished wafer is viewed with the naked eye.

A.1.3 Photolithography

Mass production with economy of scale is the primary reason for the tremendous impact VLSI has had on our society. The surface patterns of the various integrated-circuit components can be defined repeatedly using photolithography. The sequence of photolithographic steps is as illustrated in Fig. A.1.

The wafer surface is coated with a photosensitive layer called photoresist, using a spin-on technique. After this, a photographic plate with drawn patterns (e.g., a quartz plate with chromium layer for patterning) will be used to selectively expose the photoresist under a deep ultraviolet illumination (UV). The exposed areas will become softened (for positive photoresist). The exposed layer can then be removed using a chemical developer, causing the mask pattern to be duplicated on the wafer. Very fine surface geometries can be reproduced accurately by this technique. Furthermore, the patterns can be projected directly onto the wafer, or by using a separate photomask produced by a 10x "step and repeat" reduction technique as shown in Fig. A.2.

The patterned photoresist layer can be used as an effective masking layer to protect materials below from wet chemical **etching** or **reactive ion etching** (RIE). Silicon dioxide, silicon nitride, polysilicon, and metal layers can be selectively removed using the appropriate etching methods (see next section). After the etching step(s), the photoresist is stripped away, leaving behind a permanent pattern of the photomask on the wafer surface.

To make this process even more challenging, multiple masking layers (which can number more than 20 in advanced VLSI fabrication processes) must be aligned precisely on top of

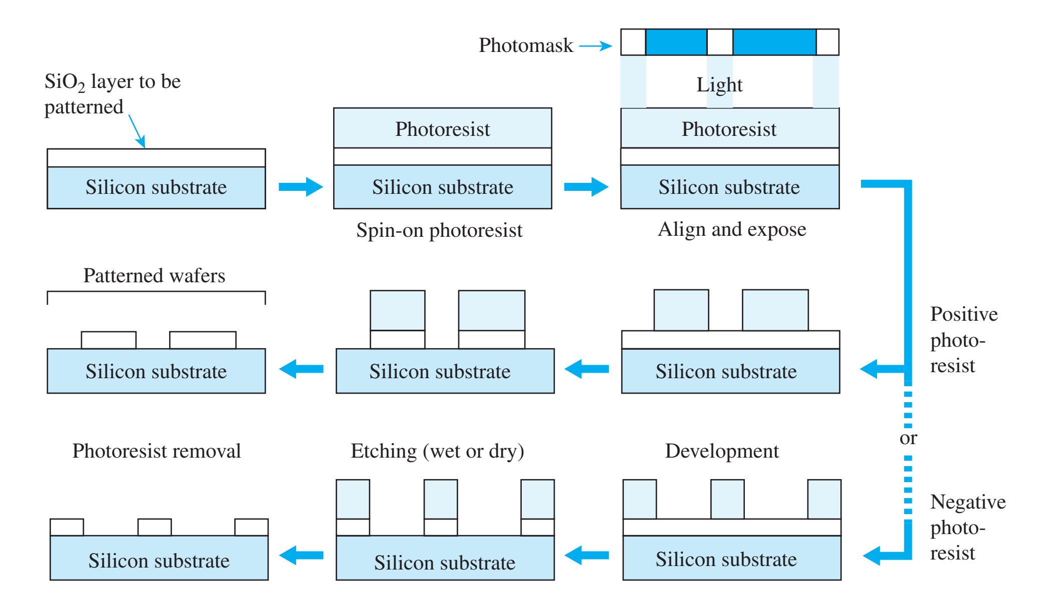


Figure A.1 Photolithography using positive or negative photoresist.

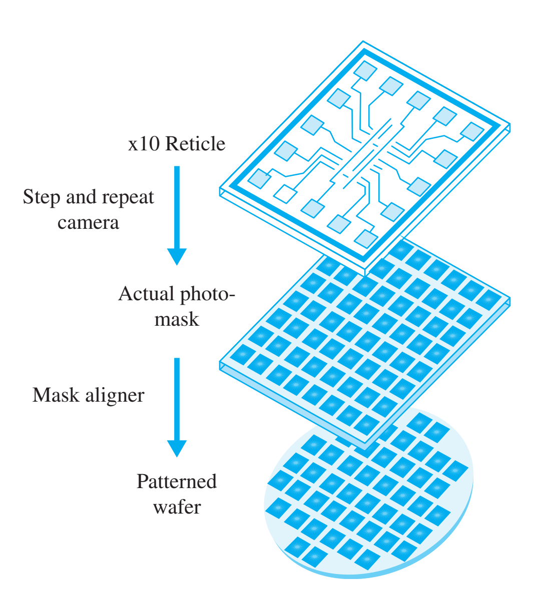


Figure A.2 Conceptual illustration of a step-and-repeat reduction technique to facilitate the mass production of integrated circuits.

previous layers. This must be done with even finer precision than the minimum geometry size of the masking patterns. This requirement imposes very critical mechanical and optical constraints on the photolithography equipment.

A.1.4 Etching

To permanently imprint the photographic patterns onto the wafer, chemical (**wet**) etching or RIE **dry etching** procedures can be used. Chemical etching is usually referred to as **wet etching**. Different chemical solutions can be used to remove different layers. For example, hydrofluoric (HF) acid can be used to etch SiO_2, potassium hydroxide (KOH) for silicon, phosphoric acid for aluminum, and so on. In wet etching, the chemical usually attacks the exposed regions that are not protected by the photoresist layer in all directions (**isotropic etching**). Depending on the thickness of the layer to be etched, a certain amount of undercut will occur. Therefore, the dimension of the actual pattern will differ slightly from the original pattern. If exact dimension is critical, RIE **dry etching** can be used. This method is essentially a directional bombardment of the exposed surface using a corrosive gas (or ions). The cross section of the etched layer is usually highly directional (**anisotropic etching**) and has the same dimension as the photoresist pattern. A comparison between isotropic and anisotropic etching is given in Fig. A.3.

A.1.5 Diffusion

Diffusion is a process by which atoms move from a high-concentration region to a low-concentration region. This is very much like a drop of ink dispersing through a glass of water except that it occurs much more slowly in solids. In VLSI fabrication, this is a method to introduce impurity atoms (dopants) into silicon to change its resistivity. The rate at which dopants diffuse in silicon is a strong function of temperature. Diffusion of impurities is usually carried out at high temperatures (1000–1200°C) to obtain the desired doping profile. When the wafer is cooled to room temperature, the impurities are essentially “frozen” in position.

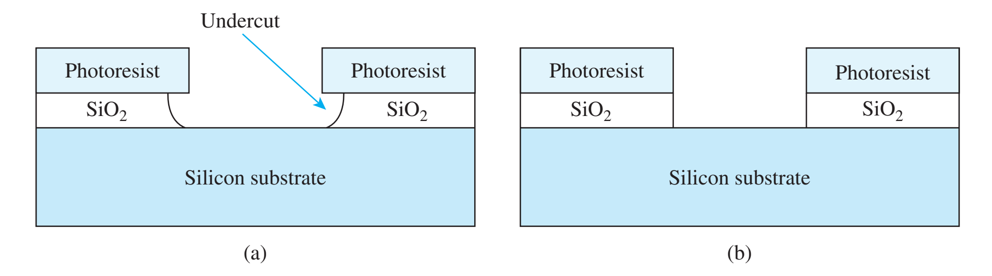


Figure A.3 (**a**) Cross-sectional view of an isotropic oxide etch with severe undercut beneath the photoresist layer. (**b**) Anisotropic etching, which usually produces a cross section with no undercut.

The diffusion process is performed in furnaces similar to those used for oxidation. The depth to which the impurities diffuse depends on both the temperature and the processing time.

The most common impurities used as **dopants** are boron, phosphorus, and arsenic. Boron is a *p*-type dopant, while phosphorus and arsenic are *n*-type dopants. These dopants can be effectively masked by thin silicon dioxide layers. By diffusing boron into an *n*-type substrate, a *pn* junction is formed (diode). If the doping concentration is heavy, the diffused layer can also be used as a conducting layer with very low resistivity.

A.1.6 Ion Implantation

Ion implantation is another method used to introduce impurities into the semiconductor crystal. An ion implanter produces ions of the desired dopant, accelerates them by an electric field, and allows them to strike the semiconductor surface. The ions become embedded in the crystal lattice. The depth of penetration is related to the energy of the ion beam, which can be controlled by the accelerating-field voltage. The quantity of ions implanted can be controlled by varying the beam current (flow of ions). Since both voltage and current can be accurately measured and controlled, ion implantation results in impurity profiles that are much more accurate and reproducible than can be obtained by diffusion. In addition, ion implantation can be performed at room temperature. Ion implantation normally is used when accurate control of the doping profile is essential for device operation.

A.1.7 Chemical Vapor Deposition

Chemical vapor deposition (CVD) is a process by which gases or vapors are chemically reacted, leading to the formation of solids on a substrate. CVD can be used to deposit various materials on a silicon substrate including SiO_2, Si_3N_4, polysilicon, and so on. For instance, if silane gas and oxygen are allowed to react above a silicon substrate, the end product, silicon dioxide, will be deposited as a solid film on the silicon wafer surface. The properties of the CVD oxide layer are not as good as those of a thermally grown oxide, but they are sufficient to allow the layer to act as an electrical insulator. The advantage of a CVD layer is that the oxide deposits at a faster rate and a lower temperature (below 500°C).

If silane gas alone is used, then a silicon layer will be deposited on the wafer. If the reaction temperature is high enough (above 1000°C), the layer deposited will be a crystalline layer (assuming that there is an exposed crystalline silicon substrate). Such a layer is called an **epitaxial** layer, and the deposition process is referred to as **epitaxy** instead of CVD. At lower temperatures, or if the substrate surface is not single-crystal silicon, the atoms will not be able to aligned along the same crystalline direction. Such a layer is called polycrystalline

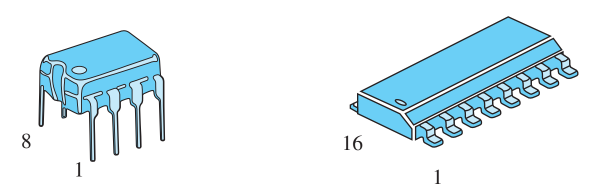


Figure A.4 Examples of an 8-pin plastic dual-in-line IC package and a 16-pin surface-mount package.

silicon (**poly Si**), since it consists of many small crystals of silicon aligned in random fashion. Polysilicon layers are normally doped very heavily to form highly conductive regions that can be used for electrical interconnections.

A.1.8 Metallization

The purpose of metallization is to interconnect the various components (transistors, capacitors, etc.) to form the desired integrated circuit. Metallization involves the deposition of a metal over the entire surface of the silicon. The required interconnection pattern is then selectively etched. The metal layer is normally deposited via a sputtering process. A pure metal disk (e.g., 99.99% aluminum target) is placed under an Ar (argon) ion gun inside a vacuum chamber. The wafers are also mounted inside the chamber above the target. The Ar ions will not react with the metal, since argon is a noble gas. However, the ions are made to physically bombard the target and literally knock metal atoms out of the target. These metal atoms will then coat all the surface inside the chamber, including the wafers. The thickness of the metal film can be controlled by the length of the sputtering time, which is normally in the range of 1 to 2 minutes. The metal interconnects can then be defined using photolithography and etching steps.

A.1.9 Packaging

A finished silicon wafer may contain several hundreds of finished circuits or chips. A chip may contain from 10 to more than 10^8 transistors; each chip is rectangular and can be up to tens of millimeters on a side. The circuits are first tested electrically (while still in wafer form) using an automatic probing station. Bad circuits are marked for later identification. The circuits are then separated from each other (by dicing), and the good circuits (dies) are mounted in packages (headers). Examples of such IC packages are given in Fig. A.4. Fine gold wires are normally used to interconnect the pins of the package to the metallization pattern on the die. Finally, the package is sealed using plastic or epoxy under vacuum or in an inert atmosphere.

A.2 VLSI Processes

Integrated-circuit fabrication technology was originally dominated by bipolar technology. By the late 1970s, metal oxide semiconductor (MOS) technology became more promising for VLSI implementation with higher packing density and lower power consumption. Since the early 1980s, complementary MOS (CMOS) technology has almost completely dominated the VLSI scene, leaving bipolar technology to fill specialized functions such as high-speed analog and RF circuits. CMOS technologies continue to evolve, and in the late 1980s, the incorporation of bipolar devices led to the emergence of high-performance bipolar-CMOS

(BiCMOS) fabrication processes that provided the best of both technologies. However, BiCMOS processes are often very complicated and costly, since they require upward of 15 to 20 masking levels per implementation—standard CMOS processes by comparison require only 10 to 12 masking levels.

The performance of CMOS and BiCMOS processes continues to improve with finer lithography resolution. However, fundamental limitations on processing techniques and semiconductor properties have prompted the need to explore alternate materials. Newly emerged SiGe and strained-Si technologies are good compromises to improve performance while maintaining manufacturing compatibility (hence low cost) with existing silicon-based CMOS fabrication equipment.

In the subsection that follows, we will examine a typical CMOS process flow, the performance of the available components, and the inclusion of bipolar devices to form a BiCMOS process.

A.2.1 Twin-Well CMOS Process

Depending on the choice of starting material (substrate), CMOS processes can be identified as ***n*-well**, ***p*-well**, or **twin-well** processes. The latter is the most complicated but most flexible in the optimization of both the *n*- and *p*-channel MOSFETs. In addition, many advanced CMOS processes may make use of trench isolation and silicon-on-insulator (SOI) technology to reduce parasitic capacitance (hence higher speed) and to improve packing density.

A modern twin-well CMOS process flow is shown in Fig. A.5. A minimum of 10 masking layers is required. In practice, most CMOS processes will also require additional layers such as *n*- and *p*-guards for better latchup immunity, a second polysilicon layer for capacitors, and multilayer metals for high-density interconnections. The inclusion of these layers would increase the total number of 15 to 20 masking layers.

The starting material for the twin-well CMOS is a *p*-type substrate. The process begins with the formation of the *p*-well and the *n*-well (Fig. A.5a). The *n*-well is required wherever *p*-channel MOSFETs are to be placed, while the *p*-well is used to house the *n*-channel MOSFETs. The well-formation procedures are similar. A thick photoresist layer is etched to expose the regions for *n*-well diffusion. The unexposed regions will be protected from the *n*-type phosphorus impurity. Phosphorus implantation is usually used for deep diffusions, since it has a large diffusion coefficient and can diffuse faster than arsenic into the substrate.

The second step is to define the active regions (region where transistors are to be placed) using a technique called **shallow trench isolation** (STI). To reduce the chance of unwanted latchup (a serious issue in CMOS technology), dry etching is used to produce trenches approximately 0.3 μm deep on the silicon surface. These trenches are then refilled using CVD oxide, followed by a planarization procedure. This results in a cross section with flat surface topology (Fig. A.5b). An alternate isolation technique is called **local oxidation of silicon** (LOCOS). This older technology uses silicon nitride (Si_3N_4) patterns to protect the penetration of oxygen during oxidation. This allows selective regions of the wafer surface to be oxidized. After a long wet-oxidation step, thick field oxide will appear in regions between transistors. This effectively produces an effect similar to that obtained in the STI process, but at the expense of large area overhead.

The next step is the formation of the polysilicon gate (Fig. A.5c). This is one of the most critical steps in the CMOS process. The thin oxide layer in the active region is first removed using wet etching followed by the growth of a high-quality thin gate oxide. Current deep-submicron CMOS processes routinely make used of oxide thicknesses as thin as 20 Å to 50 Å (1 angstrom = 10^{-8} cm). A polysilicon layer, usually arsenic doped (*n*-type), is then

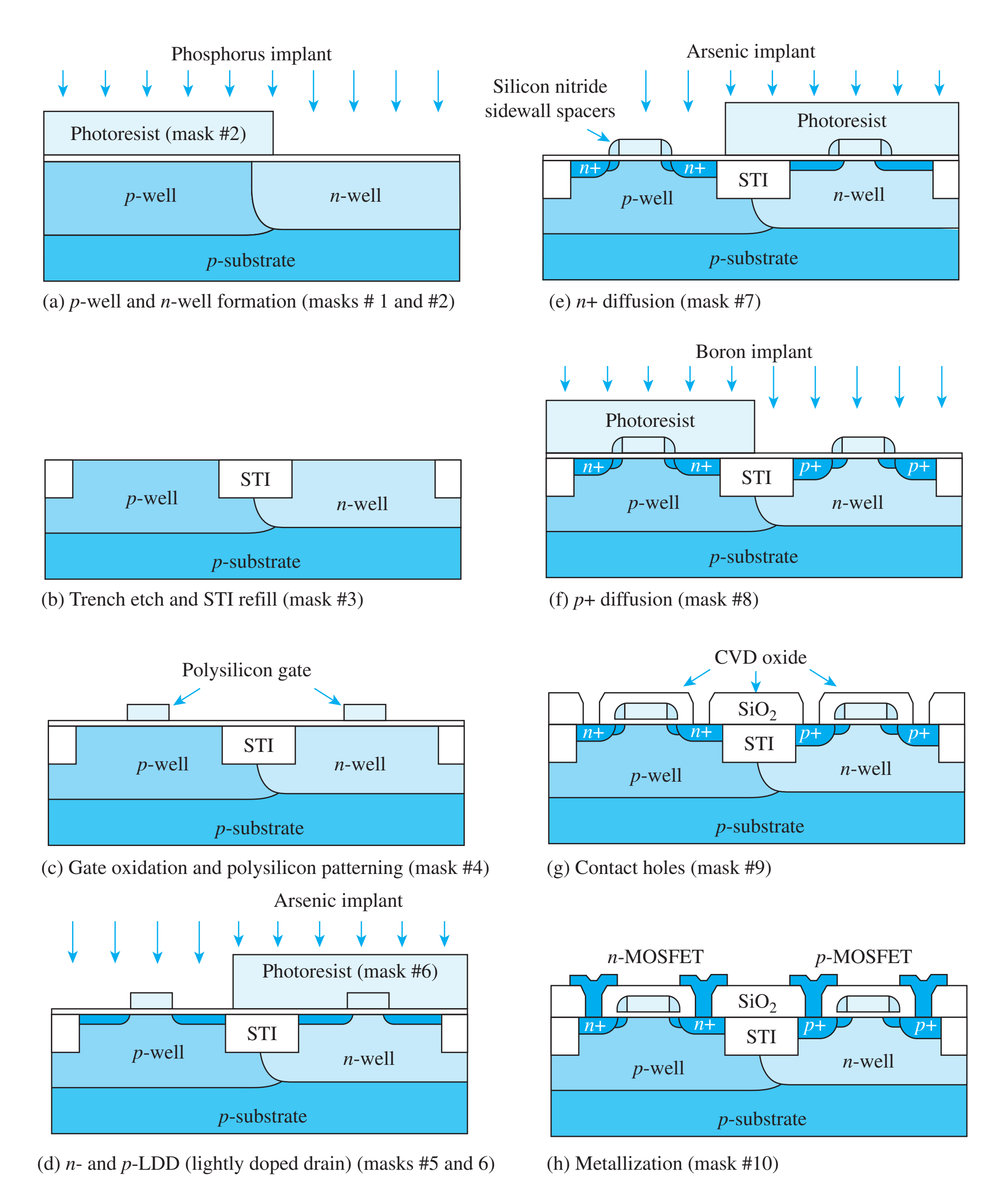


Figure A.5 A modern twin-well CMOS process flow with shallow trench isolation (STI).

deposited and patterned. The photolithography is most demanding in this step since the finest resolution is required to produce the shortest possible MOS channel length.

The polysilicon gate is a self-aligned structure and is preferred over the older type of metal gate structure. This is normally accompanied by the formation of **lightly doped drain** (LDD) regions for MOSFETs of both types to suppress the generation of **hot electrons** that might

affect the reliability of the transistors. A noncritical mask, together with the polysilicon gates, is used to form the self-aligned LDD regions (Fig. A.5d).

Prior to the $n+$ and $p+$ drain region implant, a sidewall spacer step is performed. A thick layer of silicon nitride is deposited uniformly on the wafer. Due to the conformal nature of the deposition, the thickness of the silicon nitride layer at all layer edges (i.e., at both ends of the polysilicon gate electrode) will be thicker than those deposited over a flat surface. After a timed RIE dry etch to remove all the silicon nitride layer, pockets of silicon nitride will remain at the edge of the polysilicon gate electrode (Fig. A.5e). Such pockets of silicon nitride are called sidewall spacers. They are used to block subsequent $n+$ or $p+$ source/drain implants, protecting the LDD regions.

A heavy arsenic implant can be used to form the $n+$ source and drain regions of the n-MOSFETs. The polysilicon gate also acts as a barrier for this implant to protect the channel region. A layer of photoresist can be used to block the regions where p-MOSFETs are to be formed (Fig. A.5e). The thick field oxide stops the implant and prevents $n+$ regions from forming outside the active regions. A reversed photolithography step can be used to protect the n-MOSFETs during the $p+$ boron source and drain implant for the p-MOSFETs (Fig. A.5f). Note that in both cases the separation between the source and drain diffusions—channel length—is defined by the polysilicon gate mask alone, hence the self-aligned property.

Before contact holes are opened, a thick layer of CVD oxide is deposited over the entire wafer. A photomask is used to define the contact window opening (Fig. A.5g), followed by a wet or dry oxide etch. A thin aluminum layer is then evaporated or sputtered onto the wafer. A final masking and etching step is used to pattern the interconnection (Fig. A.5h).

Not shown in the process flow is the final passivation step prior to packaging and wire bonding. A thick CVD oxide or pyrox glass is usually deposited on the wafer to serve as a protective layer.

A.2.2 Integrated Devices

Besides the obvious n- and p-channel MOSFETs, other devices can be obtained by appropriate masking patterns. These include pn junction diodes, MOS capacitors, and resistors.

A.2.3 MOSFETs

The n-channel MOSFET is the preferred device in comparison to the p-MOSFET (Fig. A.6). The electron surface mobility is two to three times higher than that for holes. Therefore, with the same device size (W and L), the n-MOSFET offers higher current drive (or lower on-resistance) and higher transconductance.

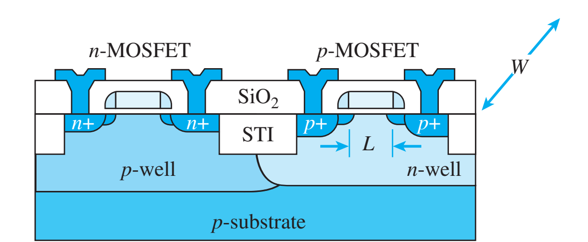


Figure A.6 Cross-sectional diagram of n- and p-MOSFETs.

In an integrated-circuit design environment, MOSFETs are characterized by their threshold voltage and by their device sizes. Usually the *n*- and *p*-channel MOSFETs are designed to have threshold voltages of similar magnitude for a particular process. The transconductance can be adjusted by changing the device surface dimensions (W and L). This feature is not available for bipolar transistor, making the design of integrated MOSFET circuits much more flexible.

A.2.4 Resistors

Resistors in integrated form are not very precise. They can be made from various diffusion regions as shown in Fig. A.7. Different diffusion regions have different resistivity. The *n* well is usually used for medium-value resistors, while the $n+$ and $p+$ diffusions are useful for low-value resistors. The actual resistance value can be defined by changing the length and width of diffused regions. The tolerance of the resistor value is very poor (20–50%), but the matching of two similar resistor values is quite good (5%). Thus circuit designers should design circuits that exploit resistor matching and should avoid designs that require a specific resistor value.

All diffused resistors are self-isolated by the reverse-biased *pn* junctions. A serious drawback for these resistors is the fact that they are accompanied by a substantial parasitic junction capacitance, making them not very useful for high-frequency applications. The reverse-biased *pn* junctions also exhibit a JFET effect, leading to a variation in the resistance value as the supply voltage is changed (a large voltage coefficient is undesirable). Since the mobilities of carriers vary with temperature, diffused resistors also exhibit a significant temperature coefficient.

A more useful resistor can be fabricated using the polysilicon layer that is placed on top of the thick field oxide. The thin polysilicon layer provides better surface area matching and hence more accurate resistor ratios. Furthermore, the polyresistor is physically separated from the substrate, resulting in a much lower parasitic capacitance and voltage coefficient.

A.2.5 Capacitors

Two types of capacitor structure are available in CMOS processes: MOS and interpoly capacitors. The latter are also similar to metal–insulator–metal (MIM) capacitors. The cross sections of these structures are as shown in Fig. A.8. The MOS gate capacitance, depicted by the center structure, is basically the gate-to-source capacitance of a MOSFET. The capacitance value is dependent on the gate area. The oxide thickness is the same as the gate oxide

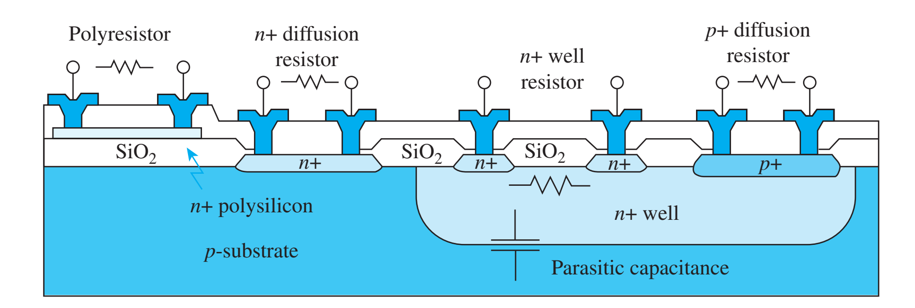


Figure A.7 Cross sections of various resistor types available from a typical *n*-well CMOS process.

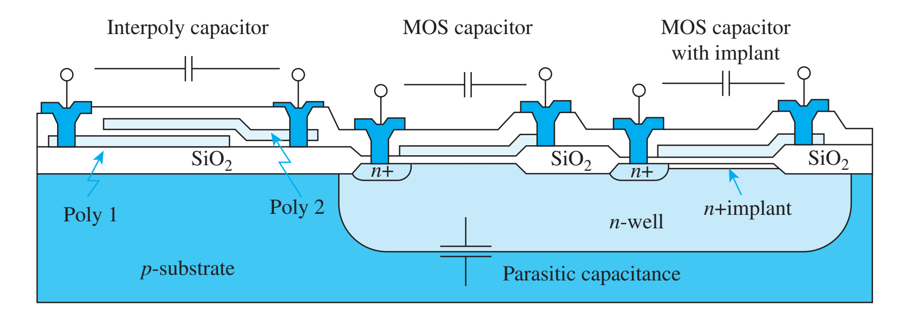


Figure A.8 Interpoly and MOS capacitors in an *n*-well CMOS process.

thickness in the MOSFETs. This capacitor exhibits a large voltage dependence. To eliminate this problem, an addition *n*+ implant is required to form the bottom plate of the capacitors, as shown in the structure on the right. Both these MOS capacitors are physically in contact with the substrate, resulting in a large parasitic *pn* junction capacitance at the bottom plate.

The interpoly capacitor exhibits near-ideal characteristics but at the expense of the inclusion of a second polysilicon layer to the CMOS process. Since this capacitor is placed on top of the thick field oxide, parasitic effects are kept to a minimum.

A third and less often used capacitor is the junction capacitor. Any *pn* junction under reversed bias produces a depletion region that acts as a dielectric between the *p* and the *n* regions. The capacitance is determined by geometry and doping levels and has a large voltage coefficient. This type of capacitor is often used as a variactor (variable capacitor) for tuning circuits. However, this capacitor works only with reverse-bias voltages.

For the interpoly and MOS capacitors, the capacitance values can be controlled to within 1%. Practical capacitance values range from 0.5 pF to a few tens of picofarads. The matching between capacitors of similar size can be within 0.1%. This property is extremely useful for designing precision analog CMOS circuits.

A.2.6 *pn* Junction Diodes

Whenever *n*-type and *p*-type diffusion regions are placed next to each other, a *pn* junction diode results. A useful structure is the *n*-well diode shown in Fig. A.9. The diode fabricated in an *n* well can provide a high breakdown voltage. This diode is essential for the input clamping circuits for protection against electrostatic discharge. The diode is also very useful as an on-chip temperature sensor by monitoring the variation of its forward voltage drop.

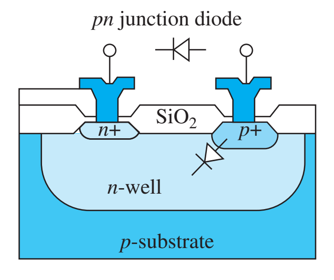


Figure A.9 A *pn* junction diode in an *n*-well CMOS process.

A.2.7 BiCMOS Process

An *npn* vertical bipolar transistor can be integrated into the *n*-well CMOS process with the addition of a *p*-base diffusion region (Fig. A.10). The characteristics of this device depend on the base width and the emitter area. The base width is determined by the difference in junction depth between the *n*+ and the *p*-base diffusions. The emitter area is determined by the junction area of the *n*+ diffusion at the emitter. The *n*-well serves as the collector for the *npn* transistor. Typically, the *npn* transistor has a β in the range of 50 to 100 and a cutoff frequency of greater than tens of gigahertz.

Normally, an *n*+ buried layer is used to reduce the series resistance of the collector, since the *n* well has a very high resistivity. However, this would further complicate the process by introducing *p*-type epitaxy and one more masking step. Other variations on the bipolar transistor includes poly-emitter and self-aligned base contact to minimize parasitic effects.

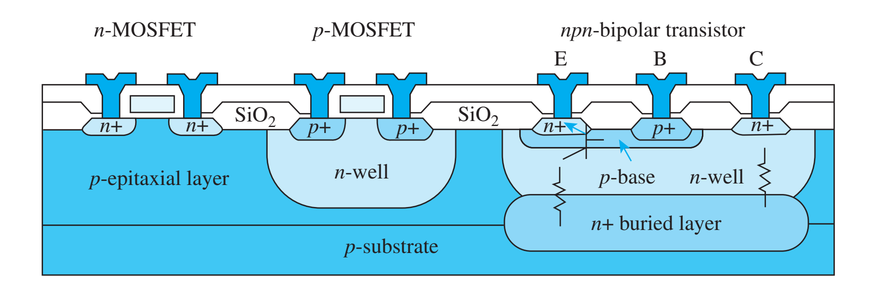


Figure A.10 Cross-sectional diagram of a BiCMOS process.

A.2.8 Lateral *pnp* Transistor

The fact that most BiCMOS processes do not have optimized *pnp* transistors makes circuit design somewhat difficult. However, in noncritical situations, a parasitic lateral *pnp* transistor can be used (Fig. A.11).

In this case, the *n* well serves as the *n*-base region, with the *p*+ diffusions as the emitter and the collector. The base width is determined by the separation between the two *p*+ diffusions. Since the doping profile is not optimized for the base–collector junctions and because the base width is limited by the minimum photolithographic resolution, the performance of this device is not very good: typically, β of around 10, and the cutoff frequency is low.

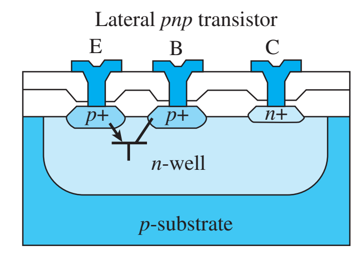


Figure A.11 Lateral *pnp* transistor.

A.2.9 *p*-Base and Pinched-Base Resistors

With the additional *p*-base diffusion in the BiCMOS process, two additional resistor structures are available. The *p*-base diffusion can be used to form a straightforward *p*-base resistor as

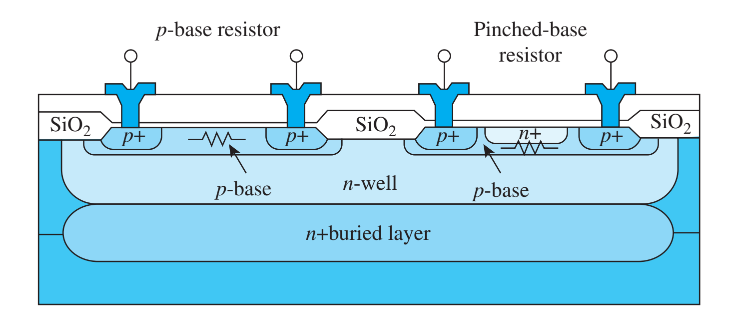


Figure A.12 *p*-base and pinched *p*-base resistors.

shown in Fig. A.12. Since the base region is usually of a relatively low doping level and has a moderate junction depth, it is suitable for medium-value resistors (a few kilohms). If a large resistor value is required, the pinched-base resistor can be used. In this structure, the *p*-base region is encroached by the *n*+ diffusion, restricting the conduction path. Resistor values in the range of 10 kΩ to 100 kΩ can be obtained. As with the diffusion resistors discussed earlier, these resistors exhibit poor tolerance and temperature coefficients but relatively good matching.

A.2.10 SiGe BiCMOS Process

With the burgeoning of wireless applications, the demand for high-performance, high-frequency RF integrated circuits is tremendous. Owing to the fundamental limitations of physical material properties, silicon-based technology was not able to compete with more expensive technologies relying on compounds from groups III through IV, such as GaAs. By incorporating a controlled amount (typically no more than 15–20% mole fraction) of germanium (Ge) into crystal silicon (Si) in the BJT's base region, the energy bandgap can be altered. The specific concentration profile of the Ge can be engineered in such a way that the energy bandgap can be gradually reduced from the pure Si region to a lower value in the SiGe region. This energy bandgap reduction produces a built-in electric field that can assist the movement of carriers, hence resulting in faster operating speed. Therefore, SiGe bipolar transistors can achieve significant higher cutoff frequency (e.g., in the 100–200 GHz range). Another benefit is that the SiGe process is compatible with existing Si-based fabrication technology, ensuring a very favorable cost/performance ratio.

To take advantage of the SiGe material characteristics, the basic bipolar transistor structure must also be modified to further reduce parasitic capacitance (for higher speed) and to improve the injection efficiency (for higher gain). A symmetric bipolar device structure is shown in Fig. A.13. The device made use of trench isolation to reduce the collector sidewall capacitance between the *n*-well/*n*+ buried layer and the *p* substrate. The emitter size and the *p*+ base contact size are defined by a self-aligned process to minimize the base–collector junction (Miller) capacitance. This type of device is called a heterojunction bipolar transistor (HBT) since the emitter–base junction is formed from two different types of material, polysilicon emitter and SiGe base. The injection efficiency is significantly better than a homojunction device (as in a conventional BJT). This advantage, coupled with the fact that base width is typically only around 50 nm, makes it easy to achieve current gain of more than 100. In addition, not shown in Fig. A.13, is the possible use of multiple layers of metallization to further reduce the device size and interconnect resistance. All these device features are necessary to complement the high-speed performance of SiGe material.

Figure A.13 Cross-sectional diagram of a symmetric self-aligned SiGe heterojunction bipolar transistor, or HBT.

A.3 VLSI Layout

The designed circuit schematic must be transformed into a layout that consists of the geometric representation of the circuit components and interconnections. Today, computer-aided design tools allow many of the conversion steps, from schematic to layout, to be carried out semi- or fully automatically. However, any good mixed-signal IC designer must have practiced full custom layout at one point or another. An example of a CMOS inverter can be used to illustrate this procedure (Fig. A.14).

The circuit must first be "flattened" and redrawn to eliminate any interconnection crossovers, similar to the requirement of a printed-circuit-board layout. Each process is made up of a specific set of masking layers. In this case, seven layers are used. Each layer is usually assigned a unique color and fill pattern for ease of identification on a computer screen or on a printed color plot. The layout begins with the placement of the transistors. For illustration purposes, the *p* and *n* MOSFETs are placed in an arrangement similar to that shown in the schematic. In practice, the designer is free to choose the most area-efficient layout. The MOSFETs are defined by the active areas overlapped by the "poly 1" layer. The MOS channel length and width are defined by the width of the "poly 1" strip and that of the active region, respectively. The *p*-MOSFET is enclosed in an *n* well. For more complex circuits, multiple *n* wells can be used for different groups of *p*-MOSFETs. The *n*-MOSFET is enclosed by the *n*+ diffusion mask to form the source and drain, while the *p*-MOSFET is enclosed by the *p*+ diffusion mask. Contact holes are placed in regions where connection to the metal layer is required. Finally, the "metal 1" layer completes the interconnections.

The corresponding cross-sectional diagram of the CMOS inverter along the AA$'$ plane is as shown in Fig. A.15. The poly-Si gates for both transistors are connected to form the input terminal, *X*. The drains of both transistors are tied together via "metal 1" to form the output terminal, *Y*. The sources of the *n*- and *p*-MOSFETs are connected to GND and V_{DD}, respectively. Note that butting contacts consist of side-by-side *n*+/*p*+ diffusions that are used to tie the body potential of the *n*- and *p*-MOSFETs to the appropriate voltage levels.

When the layout is completed, the circuit must be verified using CAD tools such as the circuit extractor, the design rule checker (DRC), and the circuit simulator. Once these verifications have been satisfied, the design can be "taped out" to a mask-making facility. A

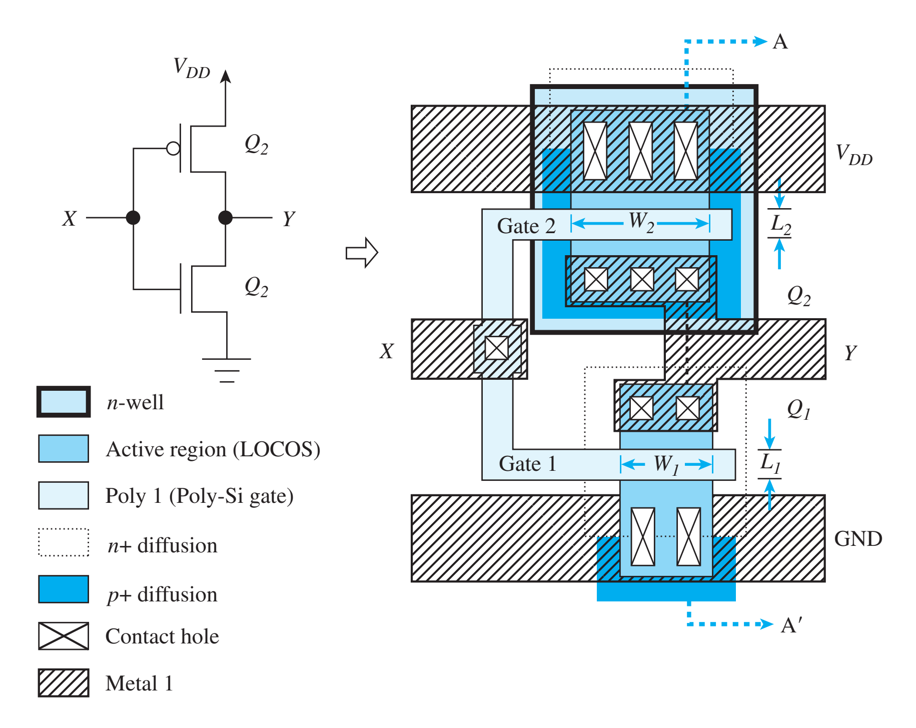


Figure A.14 A CMOS inverter schematic and its layout.

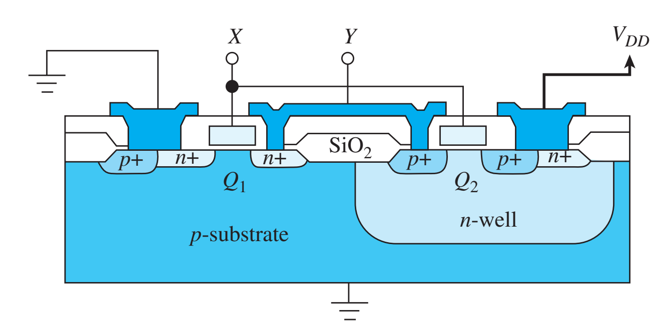


Figure A.15 Cross section along the plane AA′ of a CMOS inverter. Note that this particular layout is good for illustration purposes, but is not necessarily appropriate for latchup prevention.

pattern generator (PG) machine can then draw the geometries on a glass or quartz photoplate using electronically driven shutters. Layers are drawn one by one onto different photoplates. After these plates have been developed, clear and dark patterns resembling the geometries on the layout will result. A set of the photoplates for the CMOS inverter example is shown in Fig. A.16. Depending on whether the drawn geometries are meant to be opened as windows or kept as patterns, the plates can be **clear** or **dark field**. Note that each of these layers must be processed in sequence. The layers must be aligned within very fine tolerance to form the transistors and interconnections. Naturally, the greater the number of layers, the more difficult it is to maintain the alignment. This also requires better photolithography equipment and may

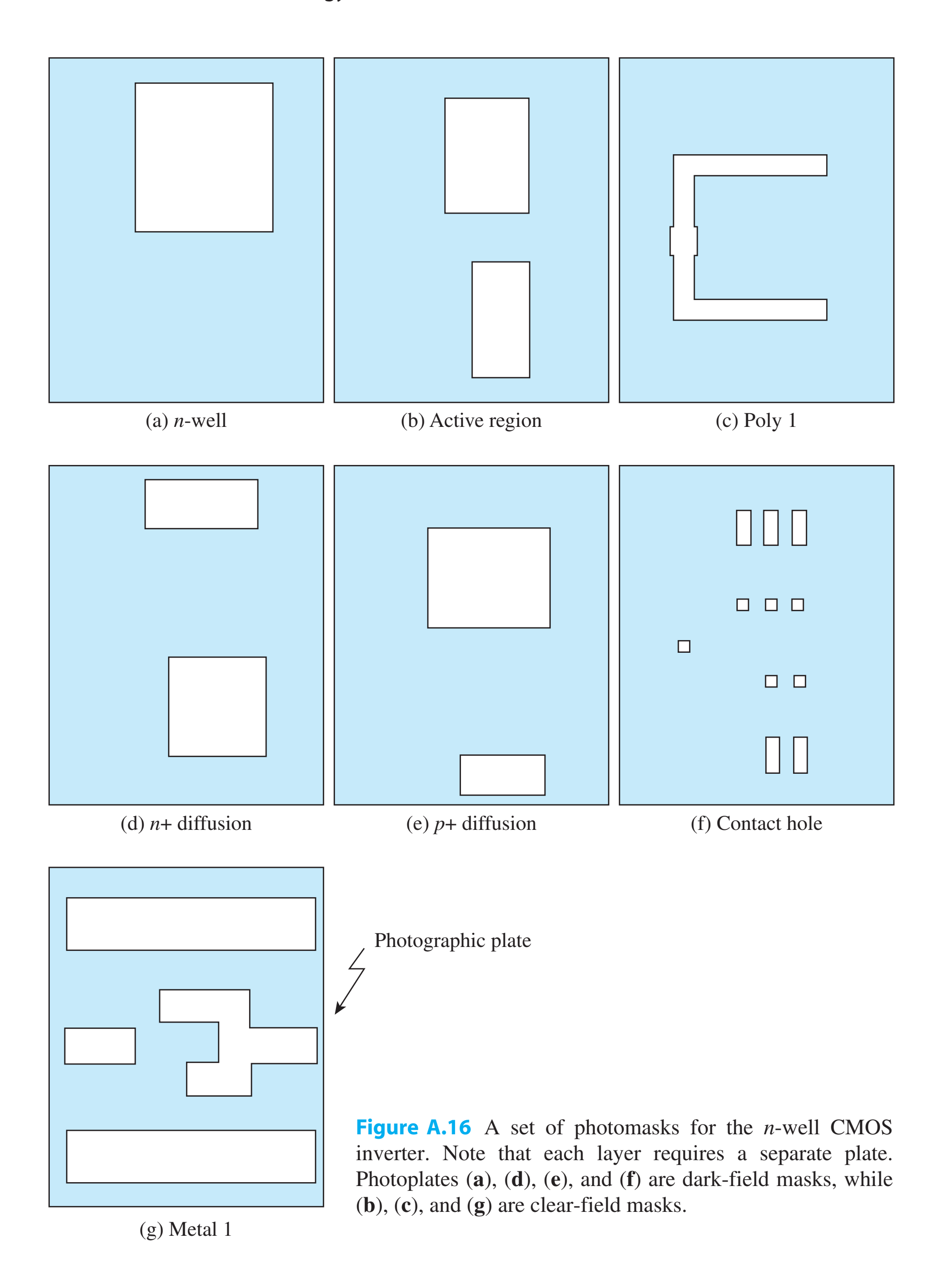


Figure A.16 A set of photomasks for the *n*-well CMOS inverter. Note that each layer requires a separate plate. Photoplates (**a**), (**d**), (**e**), and (**f**) are dark-field masks, while (**b**), (**c**), and (**g**) are clear-field masks.

result in lower yield. Hence, each additional mask will be reflected in an increase in the final cost of the IC chip.

A.4 Beyond 20 nm Technology

The rapid advancement of VLSI fabrication technology has followed a prediction called Moore's Law for more than four decades. In 1965, Gordon Moore, one of the cofounders

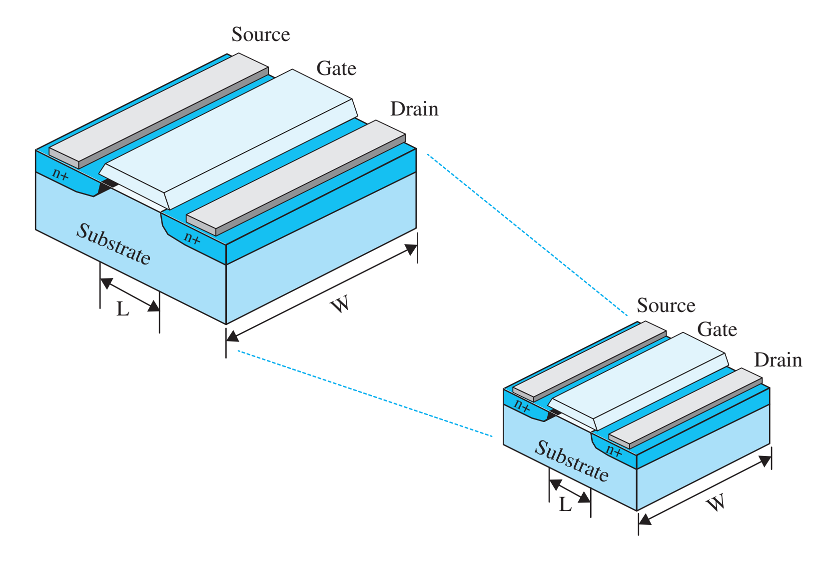


Figure A.17 MOSFET scaling consists of the reduction of both the surface and vertical dimensions. In addition, modification of the doping profiles and choice of materials are also necessary.

of Intel, foresaw that the number of transistors that can be integrated onto a VLSI chip would roughly double every two years. In order to achieve this, the size of the transistor has to be reduced accordingly. Otherwise, the size of the VLSI chip would have grown to an unacceptable size, leading to low yield and high cost. Instead of redesigning a fabrication technology from scratch every time, a scaling procedure is normally carried out. The scaling process is not only an optical shrink of the device surface layout, it also requires the reduction in vertical dimensions such as gate oxide thickness, source and drain junction depths, etc. The effect of MOSFET scaling is illustrated in Fig. A.17. VLSI fabrication technology is categorized by the minimum dimension that it can define. This is usually referred to as the channel length of the MOS gate. The reduction in device dimensions not only allows higher integration density, the shorter channel length and closer proximity of the devices also allow higher switch speed, hence better performance. As a rule of thumb, scaling cannot be carried out with an aggressive factor. Normally, 50% reducing in dimensions is achieved every two generations. Therefore, a scaling factor of approximately 0.7 is normally used. This is why we have technology nodes such as 1 μm in 1990, to 0.7 μm, 0.5 μm, 0.35 μm, 0.25 μm, 0.18 μm, 0.13 μm, 90 nm and so on.

However, this scaling approach cannot continue forever. As we approach the 20 nm technology node, the ultrathin gate dielectric and ultrashort channel length lead to an unacceptable level of gate and drain to source leakage currents. This is due to the physical limitations such as tunneling and DIBL (Drain-Induced Barrier Lowering). Some of these problems can be addressed by the use of other materials, such as high-k dielectric (e.g., HfO2) for the MOS gate.

In order to maintain the quest for an even higher level of integration, new device structures have been studied. One of the most promising technologies is the ultra-thin-body (UTB) device. In particular, the FINFET, as illustrated in Fig. A.18, has a three-dimension gate warped around a very thin slab of silicon (the fin) that stands vertically from the surface of an SOI wafer. The thin silicon fin is fully depleted during off condition to suppress drain-source leakage current. In 2014, 16 nm FINFET technology is already ready for the production of high performance VLSI chips.

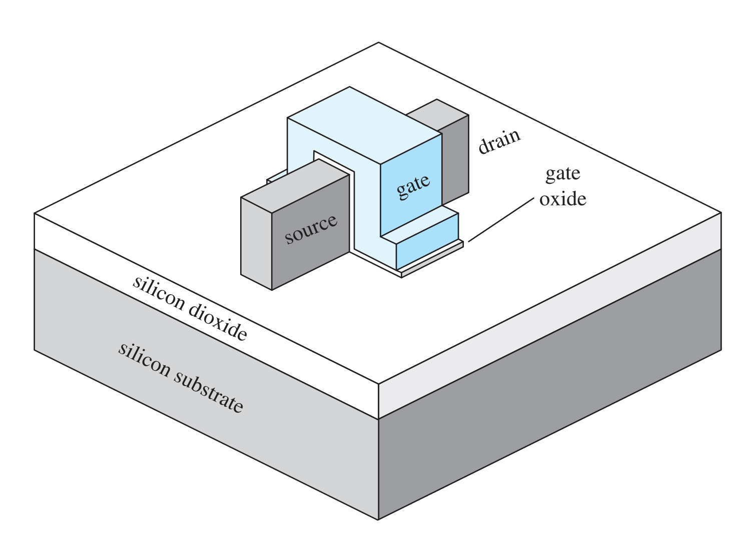


Figure A.18 A perspective view of the FINFET showing a 3D gate warped around a very thin slab of silicon fin. The source and drain contact areas are actually larger than the intrinsic device.

Summary

- This appendix presents an overview of the various aspects of VLSI fabrication procedures. This includes component characteristics, process flows, and layouts. This is by no means a complete account of state-of-the-art VLSI technologies. Interested readers should consult other references on this subject for more detailed descriptions.

Bibliography

R. S. Muller, T.I. Kamins, and M. Chan, *Device Electronics for Integrated Circuits,* 3rd ed., Hoboken, NJ, John Wiley & Sons, 2003.

J. D. Plummer, M.D. Deal, and P.B. Griffin, *Silicon VLSI Technology*, Upper Saddle River, NJ, Prentice Hall, 2000.

S. Wolf, *Microchip Manufacturing*, Lattice Press (www.latticepress.com), 2004.

APPENDIX B

SPICE DEVICE MODELS AND DESIGN SIMULATION EXAMPLES USING PSPICE AND MULTISIM

Introduction

This appendix is concerned with the very important topic of using PSpice and Multisim to simulate the operation of electronic circuits. The need for and the role of computer simulation in circuit design was described in the preface. The appendix has three sections: Section B.1 presents a brief description of the models that SPICE uses to describe the operation of op amps, diodes, MOSFETs, and BJTs. Section B.2 presents design and simulation examples using PSpice. Finally, design and simulation examples utilizing Multisim are presented in Section B.3. The examples are keyed to the book chapters and are numbered in a way that makes this relationship transparent. Thus, Example PS.2.1 refers to the first PSpice simulation example on Chapter 2 material. Similarly, Example MS.5.2 refers to the second Multisim simulation example on Chapter 5 material.

Besides the descriptions presented in this appendix, the reader will find the complete simulation files for each example on the book website.

B.1 SPICE Device Models

To the designer, the value of simulation results is a direct function of the quality of the models used for the devices. The more faithfully the model represents the various characteristics of the device, the more accurately the simulation results will describe the operation of an actual fabricated circuit. In other words, to see the effect on circuit performance of various imperfections in device operation, these imperfections must be included in the device model used by the circuit simulator.

B.1.1 The Op-Amp Model

In simulating circuits that use one or more op amps, it is useful to utilize a **macromodel** to represent each op amp. A macromodel is based on the observed terminal characteristics of the op amp rather than on the modeling of every transistor in the op-amp internal circuit. Macromodels can be developed from data-sheet specifications without knowledge of the details of the internal circuitry of the op amp.

Linear Macromodel The schematic capture of a linear macromodel for an internally compensated op amp with finite gain and bandwidth is shown in Fig. B.1. In this equivalent-circuit model, the gain constant A_{0d} of the voltage-controlled voltage source E_d corresponds to the differential gain of the op amps at dc. Resistor R_b and capacitor C_b form a

single-time-constant (STC) filter with a corner frequency

$$f_b = \frac{1}{2\pi R_b C_b} \tag{B.1}$$

The low-pass response of this filter is used to model the frequency response of the internally compensated op amp. The values of R_b and C_b used in the macromodel are chosen such that f_b corresponds to the 3-dB frequency of the op amp being modeled. This is done by arbitrarily selecting a value for either R_b or C_b (the selected value does not need to be a practical one) and then using Eq. (B.1) to compute the other value. In Fig. B.1, the voltage-controlled voltage source E_b with a gain constant of unity is used as a buffer to isolate the low-pass filter from any load at the op-amp output. Thus any op-amp loading will not affect the frequency response of the filter and hence that of the op amp.

The linear macromodel in Fig. B.1 can be further expanded to account for other op-amp nonidealities. For example, the equivalent-circuit model in Fig. B.2 can be used to model an internally compensated op amp while accounting for the following op-amp nonidealities:

1. **Input Offset Voltage (V_{OS}).** The dc voltage source V_{OS} models the op-amp input offset voltage.

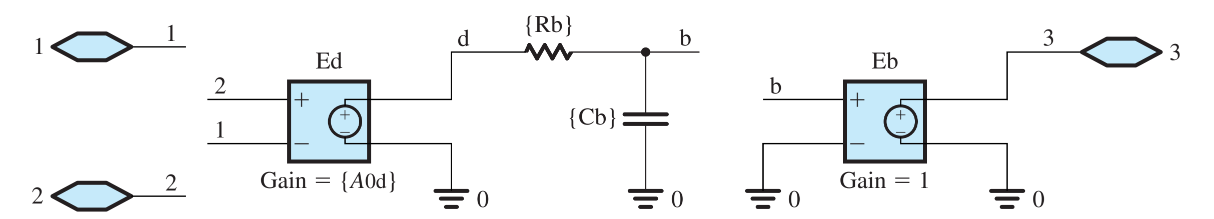


Figure B.1 A linear macromodel used to model the finite gain and bandwidth of an internally compensated op amp.

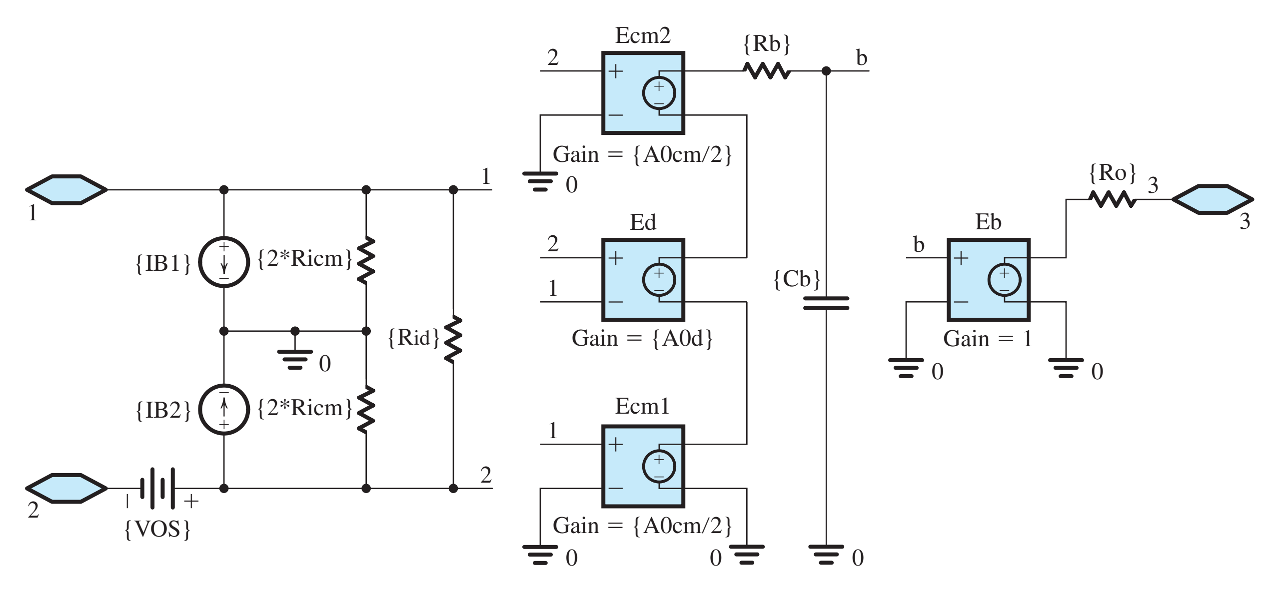


Figure B.2 A comprehensive linear macromodel of an internally compensated op amp.

2. **Input Bias Current (I_B) and Input Offset Current (I_{OS}).** The dc current sources I_{B1} and I_{B2} model the input bias current at each input terminal of the op amp, with

$$I_{B1} = I_B + \frac{I_{OS}}{2} \qquad \text{and} \qquad I_{B2} = I_B - \frac{I_{OS}}{2}$$

where I_B and I_{OS} are, respectively, the input bias current and the input offset current specified by the op-amp manufacturer.

3. **Common-Mode Input Resistance (R_{icm}).** If the two input terminals of an op amp are tied together and the input resistance (to ground) is measured, the result is the common-mode input resistance R_{icm}. In the macromodel of Fig. B.2, we have split R_{icm} into two equal parts ($2R_{icm}$), each connected between one of the input terminals and ground.

4. **Differential-Input Resistance (R_{id}).** The resistance seen between the two input terminals of an op amp is the differential input resistance R_{id}.

5. **Differential Gain at DC (A_{0d}) and Common-Mode Rejection Ratio (CMRR).** The output voltage of an op amp at dc can be expressed as

$$V_3 = A_{0d}(V_2 - V_1) + \frac{A_{0cm}}{2}(V_1 + V_2) \qquad \text{(B.2)}$$

where A_0d and A_{0cm} are, respectively, the differential and common-mode gains of the op amp at dc. For an op amp with a finite CMRR,

$$A_{0cm} = A_{0d}/\text{CMRR} \qquad \text{(B.3)}$$

where CMRR is expressed in V/V (not in dB). In the macromodel of Fig. B.2, the voltage-controlled voltage sources E_{cm1} and E_{cm2} with gain constants of $A_{0cm}/2$ account for the finite CMRR while source E_d models A_{0d}.

6. **Unity-Gain Frequency (f_t).** From Eq. (2.46), the 3-dB frequency f_b and the unity-gain frequency (or gain-bandwidth product) f_t of an internally compensated op amp with an STC frequency response are related by

$$f_b = \frac{f_t}{A_{0d}} \qquad \text{(B.4)}$$

As in Fig. B.1, the finite op-amp bandwidth is accounted for in the macromodel of Fig. B.2 by setting the corner frequency of the filter formed by resistor R_b and capacitor C_b (Eq. B.1) to equal the 3-dB frequency of the op amp, f_b.

7. **Ouput Resistance (R_o).** The resistance seen at the output terminal of an op amp is the output resistance R_o.

The linear macromodels in Figs. B.1 and B.2 assume that the op-amp circuit is operating in its linear range and do not account for its nonideal performance when large signals are present at the output. Therefore, nonlinear effects, such as output saturation and slew rate, are not modeled.

Nonlinear Macromodel The linear macromodel in Fig. B.2 can be expanded to account for the op-amp nonlinear performance. For example, the finite output voltage swing of the op amp can be modeled by placing limits on the output voltage of the voltage-controlled voltage source E_b. In PSpice, this can be done using the ETABLE component in the

analog-behavioral-modeling (ABM) library and setting the output voltage limits in the lookup table of this component. Further details on how to build nonlinear macromodels for the op amp can be found in the references on SPICE simulation. In general, robust macromodels that account for the nonlinear effects in an IC are provided by the op-amp manufacturers. Most simulators include such macromodels for some of the popular off-the-shelf ICs in their libraries. For example, PSpice and Multisim include models for the μA741, the LF411, and the LM324 op amps.

B.1.2 The Diode Model

The large-signal SPICE model for the diode is shown in Fig. B.3. The static behavior is modeled by the exponential $i-v$ relationship. Here, for generality, a constant n is included in the exponential. It is known as the **emission cofficient**, and its value ranges from 1 to 2. In our study of the diode in Chapter 4, we assumed $n=1$. The dynamic behavior is represented by the nonlinear capacitor C_D, which is the sum of the diffusion capacitance C_d and the junction capacitance C_j. The series resistance R_S represents the total resistance of the p and n regions on both sides of the junction. The value of this parasitic resistance is ideally zero, but it is typically in the range of a few ohms for small-signal diodes. For small-signal analysis, SPICE uses the diode incremental resistance r_d and the incremental values of C_d and C_j.

Table B.1 provides a partial listing of the diode-model parameters used by SPICE, all of which should be familiar to the reader. But having a good device model solves only half of the modeling problem; the other half is to determine appropriate values for the model parameters. This is by no means an easy task. The values of the model parameters are determined using a combination of characterization of the device-fabrication process and specific measurements performed on the actual manufactured devices. Semiconductor manufacturers expend enormous effort and money to extract the values of the model parameters for their devices. For discrete diodes, the values of the SPICE model parameters can be determined from the diode data sheets, supplemented if needed by key measurements. Circuit simulators (such as PSpice) include in their libraries the model parameters of some of the popular off-the-shelf components. For instance, in Example PS4.1, we will use the commercially available D1N418 pn-junction diode whose SPICE model parameters are available in PSpice.

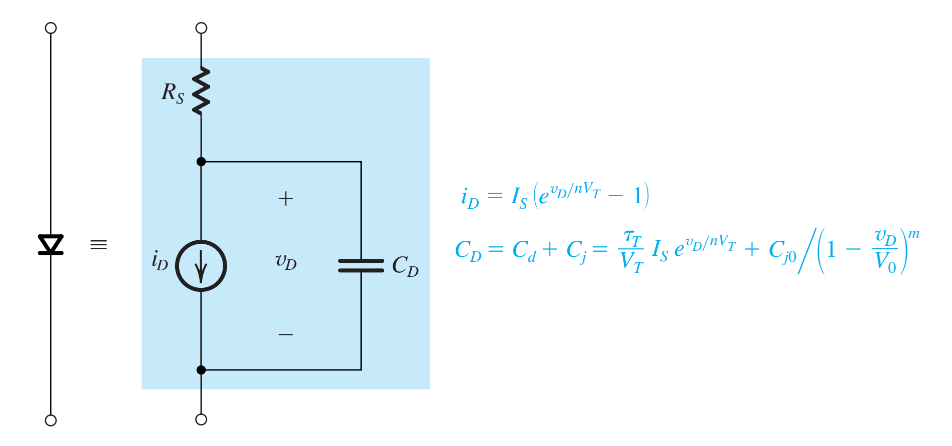


Figure B.3 The SPICE diode model.

Table B.1 Parameters of the SPICE Diode Model (Partial Listing)

SPICE Parameter	Book Symbol	Description	Units
IS	I_S	Saturation current	A
N	n	Emission coefficient	
RS	R_S	Ohmic resistance	Ω
VJ	V_0	Built-in potential	V
CJ0	C_j0	Zero-bias depletion (junction) capacitance	F
M	m	Grading coefficient	
TT	τ_T	Transit time	s
BV	V_{ZK}	Breakdown voltage	V
IBV	I_{ZK}	Reverse current at V_{ZK}	A

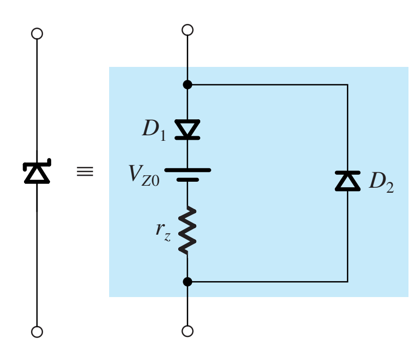


Figure B.4 Equivalent-circuit model used to simulate the zener diode in SPICE. Diode D_1 is ideal and can be approximated in SPICE by using a very small value for n (say $n = 0.01$).

B.1.3 The Zener Diode Model

The diode model in Fig. B.3 does not adequately describe the operation of the diode in the breakdown region. Hence, it does not provide a satisfactory model for zener diodes. However, the equivalent-circuit model shown in Fig B.4 can be used to simulate a zener diode in SPICE. Here, diode D_1 is an ideal diode that can be approximated in SPICE by using a very small value for n (say $n = 0.01$). Diode D_2 is a regular diode that models the forward-bias region of the zener (for most applications, the parameters of D_2 are of little consequence).

B.1.4 MOSFET Models

To simulate the operation of a MOSFET circuit, a simulator requires a mathematical model to represent the characteristics of the MOSFET. The model we derived in Chapter 5 to represent the MOSFET is a simplified or first-order model. This model, called the **square-law model** because of the quadratic $i-v$ relationship in saturation, works well for transistors with relatively *long* channels. However, for devices with *short* channels, especially deep-submicron transistors, many physical effects that we neglected come into play, with the result that the derived first-order model no longer accurately represents the actual operation of the MOSFET (see Sections 5.4.4 & 15.1.3).

The simple square-law model is useful for understanding the basic operation of the MOSFET as a circuit element and is indeed used to obtain approximate pencil-and-paper circuit designs. However, more elaborate models, which account for short-channel effects, are required to be able to predict the performance of integrated circuits with a certain degree of precision prior to fabrication. Such models have indeed been developed and continue to

be refined to more accurately represent the higher-order effects in short-channel transistors through a mix of physical relationships and empirical data. Examples include the Berkeley short-channel IGFET model (BSIM) and the EKV model, popular in Europe. Currently, semiconductor manufacturers rely on such sophisticated models to accurately represent the fabrication process. These manufacturers select a MOSFET model and then extract the values for the corresponding model parameters using both their knowledge of the details of the fabrication process and extensive measurements on a variety of fabricated MOSFETs. A great deal of effort is expended on extracting the model parameter values. Such effort pays off in fabricated circuits exhibiting performance very close to that predicted by simulation, thus reducing the need for costly redesign.

Although it is beyond the scope of this book to delve into the subject of MOSFET modeling and short-channel effects, it is important that the reader be aware of the limitations of the square-law model and of the availability of more accurate but, unfortunately, more complex MOSFET models. In fact, the power of computer simulation is more apparent when one has to use these complex device models in the analysis and design of integrated circuits.

SPICE-based simulators, like PSpice and Multisim, provide the user with a choice of MOSFET models. The corresponding SPICE model parameters (whose values are provided by the semiconductor manufacturer) include a parameter called LEVEL, which selects the MOSFET model to be used by the simulator. Although the value of this parameter is not always indicative of the accuracy, nor of the complexity of the corresponding MOSFET model, LEVEL = 1 corresponds to the simplest first-order model (called the Shichman-Hodges model), which is based on the square-law MOSFET equations presented in Chapter 5. For simplicity, we will use this model to illustrate the description of the MOSFET model parameters in SPICE and to simulate the example circuits in PSpice and Multisim. However, the reader is again reminded of the need to use a more sophisticated model than the level-1 model to accurately predict the circuit performance, especially for deep, submicron transistors.

MOSFET Model Parameters Table B.2 provides a listing of some of the MOSFET model parameters used in the level-1 model of SPICE. The reader should already be familiar with these parameters, except for a few, which are described next.

MOSFET Diode Parameters For the two reverse-biased diodes formed between each of the source and drain diffusion regions and the body (see Fig. B.4), the saturation-current density is modeled in SPICE by the parameter JS. Furthermore, based on the parameters specified in Table B.2, SPICE will calculate the depletion-layer (junction) capacitances discussed in Section 10.2.1 as

$$C_{db} = \frac{\text{CJ}}{\left(1 + \frac{V_{DB}}{\text{PB}}\right)^{\text{MJ}}}\text{AD} + \frac{\text{CJSW}}{\left(1 + \frac{V_{DB}}{\text{PB}}\right)^{\text{MJSW}}}\text{PD} \tag{B.5}$$

$$C_{sb} = \frac{\text{CJ}}{\left(1 + \frac{V_{SB}}{\text{PB}}\right)^{\text{MJ}}}\text{AS} + \frac{\text{CJSW}}{\left(1 + \frac{V_{SB}}{\text{PB}}\right)^{\text{MJSW}}}\text{PS} \tag{B.6}$$

where AD and AS are the areas, while PD and PS are the perimeters of, respectively, the drain and source regions of the MOSFET. The first capacitance term in Eqs. (B.5) and (B.6) represents the depletion-layer (junction) capacitance over the bottom plate of the drain and source regions. The second capacitance term accounts for the depletion-layer capacitance along the sidewall (periphery) of these regions. Both terms are expressed using the formula

Table B.2 Parameters of the SPICE Level-1 MOSFET Model (Partial Listing)

SPICE Parameter	Book Symbol	Description	Units
		Basic Model Parameters	
LEVEL		MOSFET model selector	
TOX	t_{ox}	Gate-oxide thickness	m
COX	C_{ox}	Gate-oxide capacitance, per unit area	F/m^2
UO	μ	Carrier mobility	$cm^2/V{\cdot}s$
KP	k'	Process transconductance parameter	A/V^2
LAMBDA	λ	Channel-length modulation coefficient	V^{-1}
		Threshold Voltage Parameters	
VTO	V_{t0}	Zero-bias threshold voltage	V
GAMMA	γ	Body-effect parameter	$V^{1/2}$
NSUB	N_A, N_D	Substrate doping	cm^{-3}
PHI	$2\phi_f$	Surface inversion potential	V
		MOSFET Diode Parameters	
JS		Body-junction saturation-current density	A/m^2
CJ		Zero-bias body-junction capacitance, per unit area over the drain/source region	F/m^2
MJ		Grading coefficient, for area component	
CJSW		Zero-bias body-junction capacitance, per unit length along F/m the sidewall (periphery) of the drain/source region	
MJSW		Grading coefficient, for sidewall component	
PB	V_0	Body-junction built-in potential	V
		MOSFET Dimension Parameters	
LD	L_{ov}	Lateral diffusion into the channel from the source/drain diffusion regions	m
WD		Sideways diffusion into the channel from the body along the width	m
		MOS Gate-Capacitance Parameters	
CGBO		Gate-body overlap capacitance, per unit channel length	F/m
CGDO	C_{ov}/W	Gate-drain overlap capacitance, per unit channel width	F/m
CGSO	C_{ov}/W	Gate-source overlap capacitance, per unit channel width	F/m

developed in Section 3.6.1 (Eq. 3.56). The values of AD, AS, PD, and PS must be specified by the user based on the dimensions of the device being used.

MOSFET Dimension and Gate-Capacitance Parameters In a fabricated MOSFET, the effective channel length L_{eff} is shorter than the nominal (or drawn) channel length L (as specified by the designer) because the source and drain diffusion regions extend slightly under the gate oxide during fabrication. Furthermore, the effective channel width W_{eff} of the MOSFET is shorter than the nominal or drawn channel width W because of the sideways diffusion into the channel from the body along the width. Based on the parameters specified in Table B.2,

$$L_{eff} = L - 2\text{LD} \tag{B.7}$$

$$W_{eff} = W - 2\text{WD} \tag{B.8}$$

In a manner analogous to using L_{ov} to denote LD, we will use the symbol W_{ov} to denote WD. Consequently, as indicated in Section 10.2.1, the gate-source capacitance C_{gs} and the gate-drain capacitance C_{gd} must be increased by an overlap component of, respectively,

$$C_{gs,ov} = W\,\text{CGSO} \tag{B.9}$$

and

$$C_{gd,ov} = W\,\text{CGDO} \tag{B.10}$$

Similarly, the gate-body capacitance C_{gb} must be increased by an overlap component of

$$C_{gb,ov} = L\,\text{CGBO} \tag{B.11}$$

The reader may have observed that there is a built-in redundancy in specifying the MOSFET model parameters in SPICE. For example, the user may specify the value of KP for a MOSFET or, alternatively, specify TOX and UO and let SPICE compute KP as UO TOX. Similarly, GAMMA can be directly specified, or the physical parameters that enable SPICE to determine it can be specified (e.g., NSUB). In any case, *the user-specified values will always take precedence over (i.e., override) those values calculated by SPICE*. As another example, note that the user has the option of either directly specifying the overlap capacitances CGBO, CGDO, and CGSO or letting SPICE compute them as CGDO = CGSO = LD COX and CGBO = WD COX.

Table B.3 provides typical values for the level-1 MOSFET model parameters of a modern 0.18-μm CMOS technology and for older 0.5-μm and 5-μm CMOS technologies. The corresponding values for the minimum channel length L_{min}, minimum channel width W_{min}, and the maximum supply voltage $(V_{DD} + |V_{SS}|)_{max}$ are as follows:

Technology	L_{min}	W_{min}	$(V_{DD} + \|V_{SS}\|)_{max}$
5-μm CMOS	5 μm	12.5 μm	10 V
0.5-μm CMOS	0.5 μm	1.25 μm	3.3 V
0.18-μm CMOS	0.18 μm	0.22 μm	1.8 V

When simulating a MOSFET circuit, the user needs to specify both the values of the model parameters and the dimensions of each MOSFET in the circuit being simulated. At least the channel length L and width W must be specified. The areas AD and AS and the perimeters PD and PS need to be specified for SPICE to model the body-junction capacitances (otherwise, zero capacitances would be assumed). The exact values of these geometry parameters depend on the actual layout of the device (Appendix A). However, to estimate these dimensions, we will assume that a metal contact is to be made to each of the source and drain regions of the MOSFET. For this purpose, typically, these diffusion regions must be extended *past* the end of the channel (i.e., in the L-direction in Fig. 5.1) by at least 2.75 L_{min}. Thus, the minimum area and perimeter of a drain/source diffusion region with a contact are, respectively,

$$\text{AD} = \text{AS} = 2.75 L_{min} W \tag{B.12}$$

Table B.3 Values of the Level-1 MOSFET Model Parameters for Two CMOS Technologies[1]

	5-μm CMOS Process		0.5-μm CMOS Process		0.18-μm CMOS Process	
	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS
LEVEL	1	1	1	1	1	1
TOX	8.50e-08	8.50e-08	9.50e-09	9.50e-09	4.08e-09	4.08e-09
UO	750	250	460	115	291	102
LAMBDA	0.01	0.03	0.1	0.2	0.08	0.11
GAMMA	1.4	0.65	0.5	0.45	0.3	0.3
VTO	1	−1	0.7	−0.8	0.5	−0.45
PHI	0.7	0.65	0.8	0.75	0.84	0.8
LD	7.00e-07	6.00e-07	8.00e-08	9.00e-08	10e-9	10e.9
JS	1.00e-06	1.00e-06	1.00e-08	5.00e-09	8.38e-6	4.00e-07
CJ	4.00e-04	1.80e-04	5.70e-04	9.30e-04	1.60e-03	1.00e-03
MJ	0.5	0.5	0.5	0.5	0.5	0.45
CJSW	8.00e-10	6.00e-10	1.20e-10	1.70e-10	2.04e-10	2.04e-10
MJSW	0.5	0.5	0.4	0.35	0.2	0.29
PB	0.7	0.7	0.9	0.9	0.9	0.9
CGBO	2.00e-10	2.00e-10	3.80e-10	3.80e-10	3.80e-10	3.50e-10
CGDO	4.00e-10	4.00e-10	4.00e-10	3.50e-10	3.67e-10	3.43e-10
CGSO	4.00e-10	4.00e-10	4.00e-10	3.50e-10	3.67e-10	3.43e-10

[1] In PSpice, we have created MOSFET parts corresponding to the above models. Readers can find these parts in the SEDRA.olb library, which is available online at www.oup.com/us/sedrasmith. The NMOS and PMOS parts for the 0.5-μm CMOS technology are labeled NMOS0P5_BODY and PMOS0P5_BODY, respectively. The NMOS and PMOS parts for the 5-μm CMOS technology are labeled NMOS5P0_BODY and PMOS5P0_BODY, respectively. Furthermore, parts NMOS5P0 and PMOS5P0 are created to correspond to, respectively, part NMOS0P5_BODY with its body connected to net 0 and part PMOS0P5_BODY with its body connected to net V_{DD}.

and

$$\text{PD} = \text{PS} = 2 \times 2.75 L_{\min} + W \tag{B.13}$$

Unless otherwise specified, we will use Eqs. (B.12) and (B.13) to estimate the dimensions of the drain/source regions in our examples.

Finally, we note that SPICE computes *the values for the parameters of the MOSFET small-signal model based on the dc operating point (bias point)*. These are then used by SPICE to perform the small-signal analysis (ac, or hand, analysis).

B.1.5 The BJT Model

SPICE uses a general form of the BJT model that we discussed in Chapter 6 (Fig. 6.5). Known as the *transport* form of the **Ebers–Moll model**, it is shown in Fig. B.5. Here, the currents of the base–emitter diode (D_{BE}) and the base–collector diode (D_{BC}) are given, respectively, by

$$i_{BE} = \frac{I_S}{\beta_F}\left(e^{v_{BE}/n_F V_T} - 1\right) \tag{B.14}$$

and

$$i_{BC} = \frac{I_S}{\beta_R}\left(e^{v_{BC}/n_R V_T} - 1\right) \tag{B.15}$$

where n_F and n_R are the emission coefficients of the BEJ and BCJ, respectively. These coefficients are generalizations of the constant n of the *pn*-junction diode (Fig. B.3).

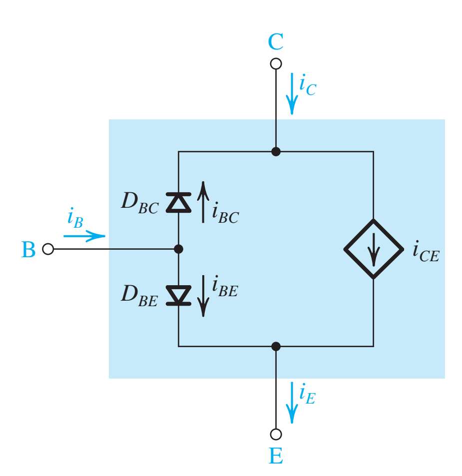


Figure B.5 The transport form of the Ebers–Moll model for an *npn* BJT.

(We have so far assumed $n_F = n_R = 1$). The parameters β_F and β_R are, respectively, the forward and reverse β of the BJT. The reverse β is the current gain obtained when the collector and emitter are interchanged and is much smaller than the forward β. In fact, $\beta_R \ll 1$. The controlled current-source i_{CE} in the transport model is defined as

$$i_{CE} = I_S\left(e^{v_{BE}/n_F V_T} - e^{v_{BC}/n_R V_T}\right) \tag{B.16}$$

Observe that i_{CE} represents the current component of i_C and i_E that arises as a result of the minority carrier diffusion across the base, or **carrier transport** across the base (hence the name transport model).

The transport model can account for the Early effect in a forward-biased BJT by including the factor $(1 - v_{BC}/V_A)$ in the expression for the transport current i_{CE} as follows:

$$i_{CE} = I_S\left(e^{v_{BE}/n_F V_T} - e^{v_{BC}/n_R V_T}\right)\left(1 - \frac{v_{BC}}{V_A}\right) \tag{B.17}$$

Figure B.6 shows the model used in SPICE. Here, resistors r_x, r_E, and r_C are added to represent the ohmic resistance of, respectively, the base, emitter, and collector regions. The dynamic operation of the BJT is modeled by two nonlinear capacitors, C_{BC} and C_{BE}. Each of these capacitors generally includes a diffusion component (i.e., C_{DC} and C_{DE}) and a depletion or junction component (i.e., C_{JC} and C_{JE}) to account for the charge-storage effects within the BJT (as described in Section 10.6.2). Furthermore, the BJT model includes a depletion junction capacitance C_{JS} to account for the collector–substrate junction in integrated-circuit BJTs, where a reverse-biased *pn* junction is formed between the collector and the substrate (which is common to all components of the IC).

For small-signal (ac) analysis, the SPICE BJT model is equivalent to the hybrid-π model of Fig. 7.24, but augmented with r_E, r_C, and (for IC BJTs) C_{JS}. Furthermore, the model includes a large resistance r_μ between the base and collector (in parallel with C_μ) to account for the dependence of i_1 on v_{CB}. The resistance r_μ is very large, typically greater than $10\beta r_o$.

Although Fig. B.5 shows the SPICE model for the *npn* BJT, the corresponding model for the *pnp* BJT can be obtained by reversing the direction of the currents and the polarity of the diodes and terminal voltages.

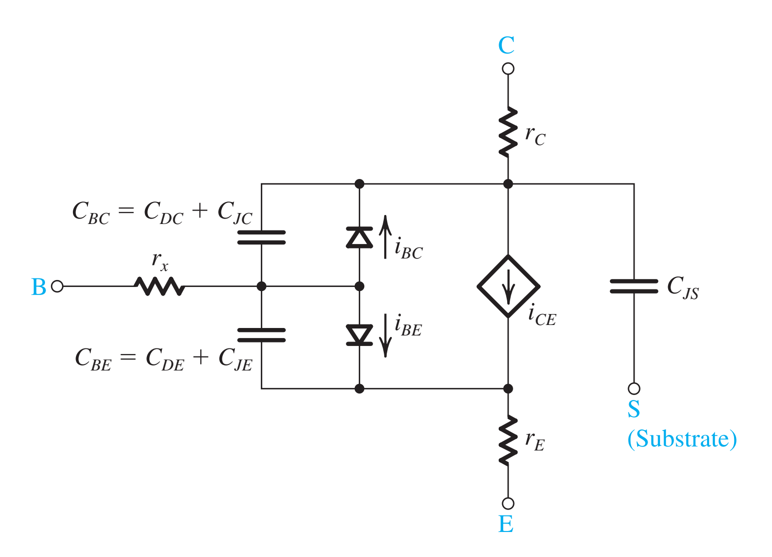


Figure B.6 The SPICE large-signal model for an *npn* BJT.

The SPICE Gummel–Poon Model of the BJT The BJT model described above lacks a representation of some second-order effects present in actual devices. One of the most important such effects is the variation of the current gains, β_F and β_R, with the current i_C. The Ebers–Moll model assumes β_F and β_R to be constant, thereby neglecting their current dependence (as depicted in Fig. 6.19). To account for this, and other second-order effects, SPICE uses a more accurate, yet more complex, BJT model called the Gummel–Poon model (named after H. K. Gummel and H. C. Poon, two pioneers in this field). This model is based on the relationship between the electrical terminal characteristics of a BJT and its base charge. It is beyond the scope of this book to delve into the model details. However, it is important for the reader to be aware of the existence of such a model.

In SPICE, the Gummel–Poon model automatically simplifies to the Ebers–Moll model when certain model parameters are not specified. Consequently, the BJT model to be used by SPICE need not be explicitly specified by the user (unlike the MOSFET case in which the model is specified by the LEVEL parameter). For discrete BJTs, the values of the SPICE model parameters can be determined from the data specified on the BJT data sheets, supplemented (if needed) by key measurements. For instance, in Example PS5.6.1, we will use the Q2N3904 *npn* BJT (from Fairchild Semiconductor) whose SPICE model is available in PSpice. In fact, the PSpice and Multisim library already includes the SPICE model parameters for many of the commercially available discrete BJTs. For IC BJTs, the values of the SPICE model parameters are determined by the IC manufacturer (using both measurements on the fabricated devices and knowledge of the details of the fabrication process) and are provided to the IC designers.

The SPICE BJT Model Parameters Table B.4 provides a listing of some of the BJT model parameters used in SPICE. The reader should be already familiar with these parameters. In the absence of a user-specified value for a particular parameter, SPICE uses a default value that typically results in the corresponding effect being ignored. For example, if no value is specified for the forward Early voltage (VAF), SPICE assumes that VAF $= \infty$

and does not account for the Early effect. Although ignoring VAF can be a serious issue in some circuits, the same is not true, for example, for the value of the reverse Early voltage (VAR).

The BJT Model Parameters BF and BR in SPICE Before leaving the SPICE model, a comment on β is in order. SPICE interprets the user-specified model parameters BF and BR as the *ideal maximum* values of the forward and reverse dc current gains, respectively, versus the operating current. These parameters are not equal to the constant-current-independent parameters $\beta_F(\beta_{dc})$ and β_R used in the Ebers–Moll model for the forward and reverse dc current gains of the BJT. SPICE uses a current-dependent model for β_F and β_R, and the user can specify other parameters (not shown in Table B.4) for this model. Only when such parameters are not specified, and the Early effect is neglected, will SPICE assume that β_F and β_R are constant and equal to BF and BR, respectively. Furthermore, SPICE computes values for both β_{dc} and β_{ac}, the two parameters that we generally assume to be approximately equal. SPICE then uses β_{ac} to perform small-signal (ac) analysis.

Table B.4 Parameters of the SPICE BJT Model (Partial Listing)

SPICE Parameter	Book Symbol	Description	Units
IS	I_S	Saturation current	A
BF	β_F	Ideal maximum forward current gain	
BR	β_R	Ideal maximum reverse current gain	
NF	n_F	Forward current emission coefficient	
NR	n_R	Reverse current emission coefficient	
VAF	V_A	Forward Early voltage	V
VAR		Reverse Early voltage	V
RB	r_x	Zero-bias base ohmic resistance	Ω
RC	r_C	Collector ohmic resistance	Ω
RE	r_E	Emitter ohmic resistance	Ω
TF	τ_F	Ideal forward transit time	s
TR	τ_R	Ideal reverse transit time	s
CJC	$C_{\mu 0}$	Zero-bias base–collector depletion (junction) capacitance	F
MJC	m_{BCJ}	Base–collector grading coefficient	
VJC	V_{0c}	Base–collector built-in potential	V
CJE	C_{je0}	Zero-bias base–emitter depletion (junction) capacitance	F
MJE	m_{BEJ}	Base–emitter grading coefficient	
VJE	V_{0e}	Base–emitter built-in potential	V
CJS		Zero-bias collector–substrate depletion (junction) capacitance	F
MJS		Collector–substrate grading coefficient	
VJS		Collector–substrate built-in potential	V

B.2 PSpice Examples

Example PS.2.1

Performance of a Noninverting Amplifier

Consider an op amp with a differential input resistance of 2 MΩ, an input offset voltage of 1 mV, a dc gain of 100 dB, and an output resistance of 75 Ω. Assume the op amp is internally compensated and has an STC frequency response with a gain–bandwidth product of 1 MHz.

(a) Create a subcircuit model for this op amp in PSpice.
(b) Using this subcircuit, simulate the closed-loop noninverting amplifier in Fig. 2.12 with resistors $R_1 =$ 1 kΩ and $R_2 = 100$ kΩ to find:
  (i) Its 3-dB bandwidth f_{3dB}.
  (ii) Its output offset voltage V_{OSout}.
  (iii) Its input resistance R_{in}.
  (iv) Its output resistance R_{out}.
(c) Simulate the step response of the closed-loop amplifier, and measure its rise time t_r. Verify that this time agrees with the 3-dB frequency measured above.

Solution

To model the op amp in PSpice, we use the equivalent circuit in Fig. B.2, but with $R_{id} = 2$ MΩ, $R_{icm} = \infty$ (open circuit), $I_{B1} = I_{B2} = 0$ (open circuit), $V_{OS} = 1$ mV, $A_{0d} = 10^5$ V/V, $A_{0cm} = 0$ (short circuit), and $R_o = 75$ Ω. Furthermore, we set $C_b = 1$ μF and $R_b = 15.915$ kΩ to achieve an $f_t = 1$ MHz.

To measure the 3-dB frequency of the closed-loop amplifier, we apply a 1-V ac voltage at its input, perform an ac-analysis simulation in PSpice, and plot its output versus frequency. The output voltage, plotted in Fig. B.7, corresponds to the gain of the amplifier because we chose an input voltage of 1 V.

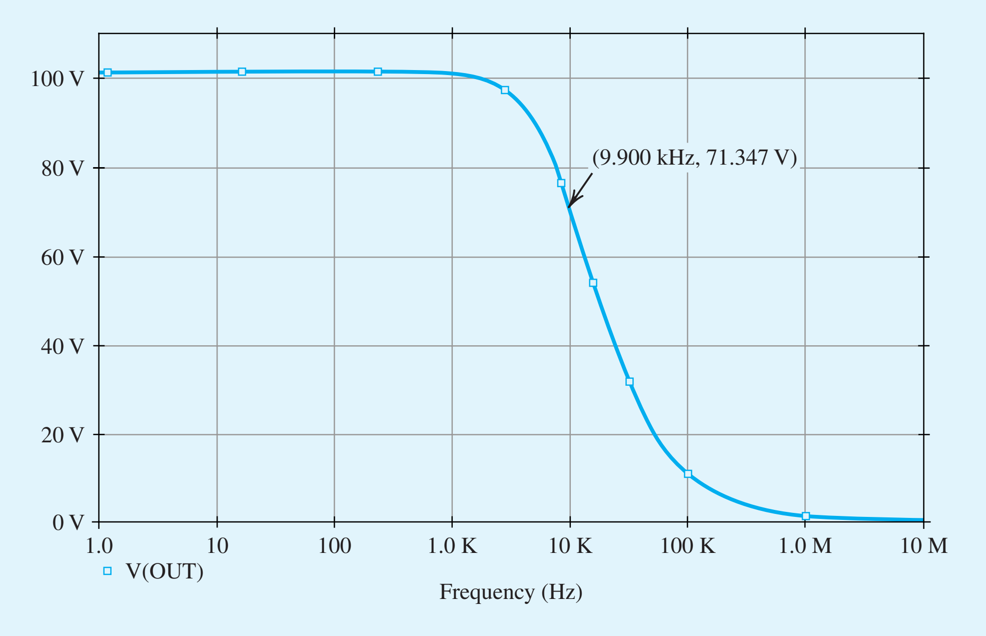


Figure B.7 Frequency response of the closed-loop amplifier in Example PS.2.1.

Example PS.2.1 *continued*

Thus, from Fig. B.7, the closed-loop amplifier has a dc gain of $G_0 = 100.9$ V/V, and the frequency at which its gain drops to $G_0/\sqrt{2} = 71.35$ V/V is $f_{3dB} = 9.9$ kHz, which agrees with Eq. (B.7).

The input resistance R_{in} corresponds to the reciprocal of the current drawn out of the 1-V ac voltage source used in the above ac-analysis simulation at 0.1 Hz. (Theoretically, R_{in} is the small-signal input resistance at dc. However, ac-analysis simulations must start at frequencies greater than zero, so we use 0.1 Hz to approximate the dc point.) Accordingly, R_{in} is found to be 2 GΩ.

To measure R_{out}, we short-circuit the amplifier input to ground, inject a 1-A ac current at its output, and perform an ac-analysis simulation. R_{out} corresponds to the amplifier output voltage at 0.1 Hz and is found to be 76 mΩ. Although an ac test voltage source could equally well have been used to measure the output resistance in this case, it is a good practice to attach a current source rather than a voltage source between the output and ground. This is because an ac current source appears as an open circuit when the simulator computes the dc bias point of the circuit while an ac voltage source appears as a short circuit, which can erroneously force the dc output voltage to zero. For similar reasons, an ac test voltage source should be attached in series with the biasing dc voltage source for measuring the input resistance of a voltage amplifier.

A careful look at R_{in} and R_{out} of the closed-loop amplifier reveals that their values have, respectively, increased and decreased by a factor of about 1000, relative to the corresponding resistances of the op amp. Such a large input resistance and small output resistance are indeed desirable characteristics for a voltage amplifier. This improvement in the small-signal resistances of the closed-loop amplifier is a direct consequence of applying negative feedback (through resistors R_1 and R_2) around the open-loop op amp. We will study negative feedback in Chapter 11, where we will also learn how the improvement factor (1000 in this case) corresponds to the ratio of the open-loop op-amp gain (10^5) to the closed-loop amplifier gain (100).

From Eqs. (2.55) and (2.53), the closed-loop amplifier has an STC low-pass response given by

$$\frac{V_o(s)}{V_i(s)} = \frac{G_0}{1 + \dfrac{s}{2\pi f_{3dB}}}$$

As described in Appendix E, the response of such an amplifier to an input step of height V_{step} is given by

$$v_O(t) = V_{final}\left(1 - e^{-t/\tau}\right) \tag{B.18}$$

where $V_{final} = G_0 V_{step}$ is the final output-voltage value (i.e., the voltage value toward which the output is heading) and $\tau = 1/\left(2\pi f_{3dB}\right)$is the time constant of the amplifier. If we define $t_{10\%}$ and $t_{90\%}$ to be the time it takes for the output waveform to rise to, respectively, 10% and 90% of V_{final}, then from Eq. (B.18), $t_{10\%} \simeq 0.1\tau$ and $t_{90\%} \simeq 2.3\tau$. Therefore, the rise time t_r of the amplifier can be expressed as

$$t_r = t_{90\%} - t_{10\%} = 2.2\tau = \frac{2.2}{2\pi f_{3dB}}$$

Therefore, if $f_{3dB} = 9.9$ kHz, then $t_r = 35.4$ µs. To simulate the step response of the closed-loop amplifier, we apply a step voltage at its input, using a piecewise-linear (PWL) source (with a very short rise time); then perform a transient-analysis simulation, and measure the voltage at the output versus time. In our simulation, we applied a 1-V step input, plotted the output waveform in Fig. B.8, and measured t_r to be 35.3 µs.

The linear macromodels in Figs. B.1 and B.2 assume that the op-amp circuit is operating in its linear range; they do not account for its nonideal performance when large signals are present at the output. Therefore, nonlinear effects, such as output saturation and slew rate, are not modeled. This is why, in the step response of Fig. B.8, we could see an output voltage of 100 V when we applied a 1-V step input. However, IC op amps are not capable of producing such large output voltages. Hence, a designer must be very careful when using these models.

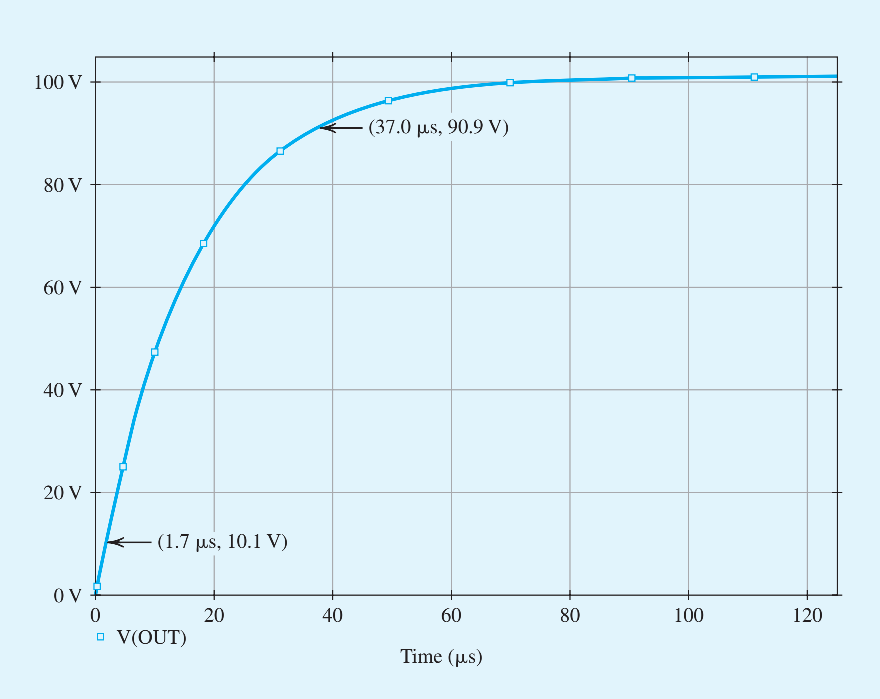


Figure B.8 Step response of the closed-loop amplifier in Example PS.2.1.

It is important to point out that we also saw output voltages of 100 V or so in the ac analysis of Fig. B.7, where for convenience we applied a 1-V ac input to measure the gain of the closed-loop amplifier. So, would we see such large output voltages if the op-amp macromodel accounted for nonlinear effects (particularly output saturation)? The answer is yes, because in an ac analysis PSpice uses a linear model for nonlinear devices with the linear-model parameters evaluated at a bias point. Thus, we must keep in mind that the voltage magnitudes encountered in an ac analysis may not be realistic. In this case, the voltage and current ratios (e.g., the output-to-input voltage ratio as a measure of voltage gain) are of importance to the designer.

Example PS.2.2

Characteristics of the 741 Op Amp

Consider the μA741 op amp whose macromodel is available in PSpice. Use PSpice to plot the open-loop gain and hence determine f_t. Also, investigate the SR limitation and the output saturation of this op amp.

Solution

Figure B.9 shows the schematic capture used to simulate the frequency response of the μA741 op amp.[1] The μA741 part has seven terminals. Terminals 7 and 4 are, respectively, the positive and negative

[1]The reader is reminded that the schematic capture diagram and the corresponding PSpice simulation files of all SPICE examples in this book can be found on the text's website (www.oup.com/us/sedrasmith).

Example PS.2.2 *continued*

Figure B.9 Simulating the frequency response of the μA741 op-amp in Example PS.2.2.

dc power-supply terminals of the op amp. The 741-type op amps are typically operated from ±15-V power supplies; therefore we connected the dc voltage sources $V_{CC} = +15$ V and $V_{EE} = -15$ V to terminals 7 and 4, respectively. Terminals 3 and 2 of the μA741 part correspond to the positive and negative input terminals, respectively, of the op amp. In general, as outlined in Section 2.1.3, the op-amp input signals are expressed as

$$v_{INP} = V_{CM} + \frac{V_d}{2}$$

$$v_{INN} = V_{CM} - \frac{V_d}{2}$$

where v_{INP} and v_{INN} are the signals at, respectively, the positive- and negative-input terminals of the op amp with V_{CM} being the common-mode input signal (which sets the dc bias voltage at the op-amp input terminals) and V_d being the differential input signal to be amplified. The dc voltage source V_{CM} in Fig. B.9 is used to set the common-mode input voltage. Typically, V_{CM} is set to the average of the dc power-supply voltages V_{CC} and V_{EE} to maximize the available input signal swing. Hence, we set $V_{CM} = 0$. The voltage source V_d in Fig. B.9 is used to generate the differential input signal V_d. This signal is applied differentially to the op-amp input terminals using the voltage-controlled voltage sources E_p and E_n, whose gain constants are set to 0.5.

Terminals 1 and 5 of part μA741 are the offset-nulling terminals of the op amp (as depicted in Fig. 2.36). However, a check of the PSpice netlist of this part (by selecting Edit → PSpice Model, in the Capture menus), reveals that these terminals are floating; therefore the offset-nulling characteristic of the op amp is not incorporated in this macromodel.

To measure f_t of the op amp, we set the voltage of source V_d to be 1-V ac, perform an ac-analysis simulation in PSpice, and plot the output voltage versus frequency as shown in Fig. B.10. Accordingly, the frequency at which the op-amp voltage gain drops to 0 dB is $f_t = 0.9$ MHz (which is close to the 1-MHz value reported in the data sheets for 741-type op amps).

To determine the slew rate of the μA741 op amp, we connect the op amp in a unity-gain configuration, as shown in Fig. B.11, apply a large pulse signal at the input with very short rise and fall times to

---

In these schematics (as shown in Fig. B.13), we use variable parameters to enter the values of the various circuit components. This allow one to investigate the effect of changing component values by simply changing the corresponding parameter values.

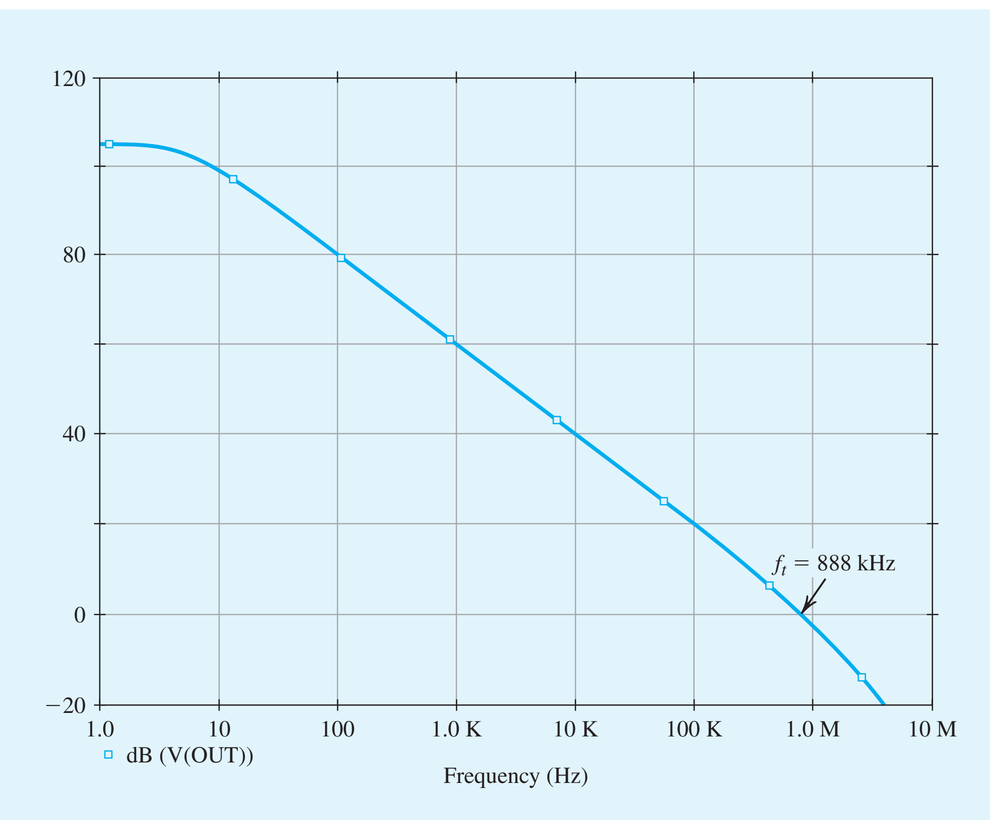


Figure B.10 Frequency response of the µA741 op amp in Example PS.2.2.

cause slew-rate limiting at the output, perform a transient-analysis simulation in PSpice, and plot the output voltage as shown in Fig. B.12. The slope of the slew-rate limited output waveform corresponds to the slew-rate of the op amp and is found to be SR = 0.5 V/µs (which agrees with the value specified in the data sheets for 741-type op amps).

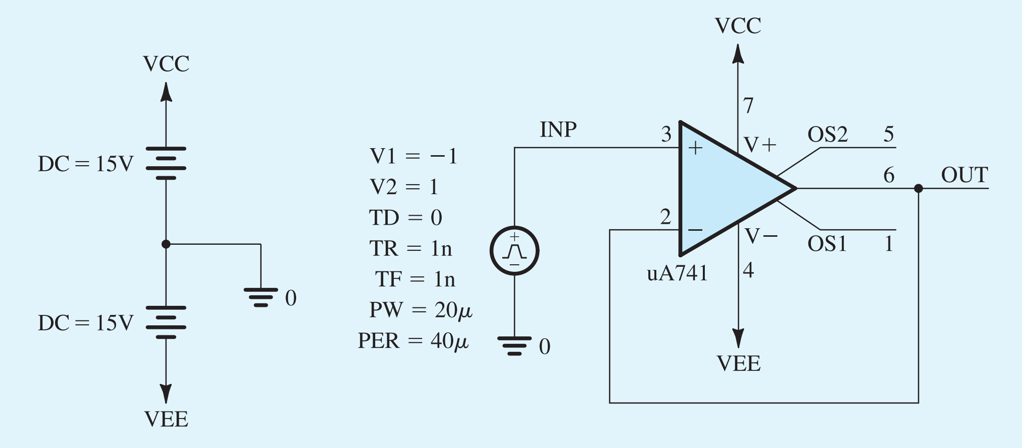


Figure B.11 Circuit for determining the slew rate of the µA741 op amp in Example PS.5.2.2.

Example PS.2.2 *continued*

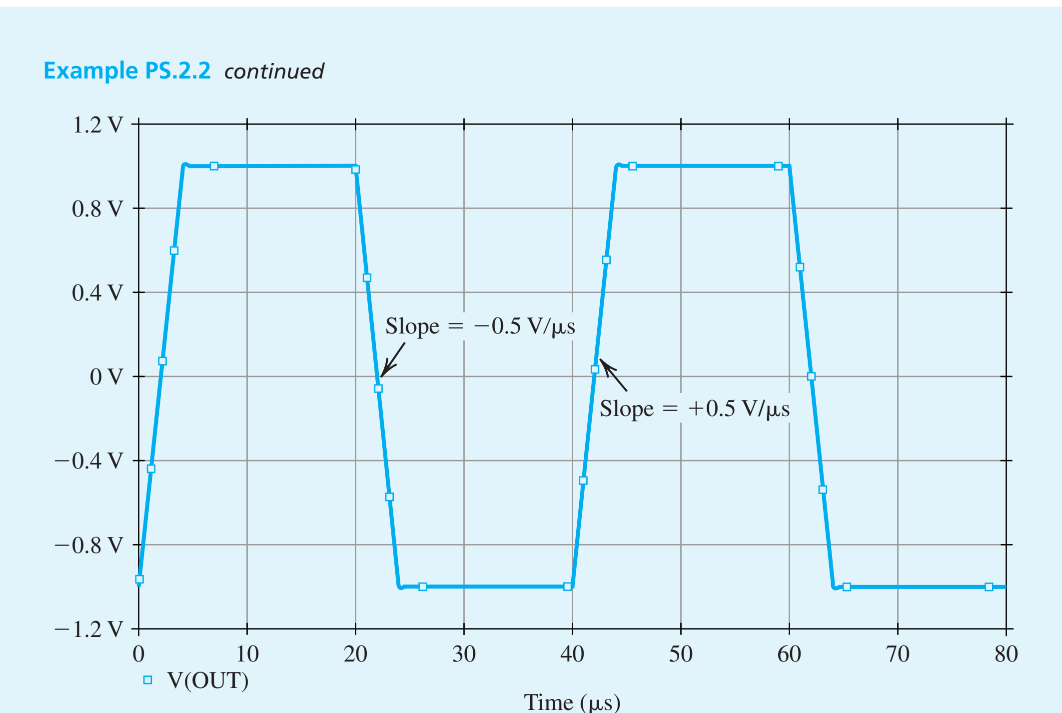


Figure B.12 Square-wave response of the μA741 op amp connected in the unity-gain configuration shown in Fig. B.11.

To determine the maximum output voltage of the μA741 op amp, we set the dc voltage of the differential voltage source V_d in Fig. B.9 to a large value, say +1 V, and perform a bias-point simulation in PSpice. The corresponding dc output voltage is the positive-output saturation voltage of the op amp. We repeat the simulation with the dc differential input voltage set to –1 V to find the negative-output saturation voltage. Accordingly, we find that the μA741 op amp has a maximum output voltage $V_{omax} = 14.8$ V.

Example PS.4.1

Design of a DC Power Supply

In this example, we will design a dc power supply using the rectifier circuit whose capture schematic is shown in Fig. B.13. This circuit consists of a full-wave diode rectifier, a filter capacitor, and a zener voltage regulator. The only perhaps puzzling component is the $R_{\text{isolation}}$, the 100-MΩ resistor between the secondary winding of the transformer and ground. This resistor is included to provide dc continuity and thus “keep SPICE happy”; it has little effect on circuit operation.

Let it be required that the power supply (in Fig. B.13) provide a nominal dc voltage of 5 V and be able to supply a load current I_{load} as large as 25 mA; that is, R_{load} can be as low as 200 Ω. The power supply is fed from a 120-V (rms) 60-Hz ac line. Note that in the PSpice schematic (Fig. B.13), we use a sinusoidal voltage source with a 169-V peak amplitude to represent the 120-V rms supply (as 120-V rms = 169-V peak). Assume the availability of a 5.1-V zener diode having $r_z = 10$ Ω at $I_Z = 20$ mA (and thus $V_{Z0} = 4.9$ V), and that the required minimum current through the zener diode is $I_{Z\text{min}} = 5$ mA.

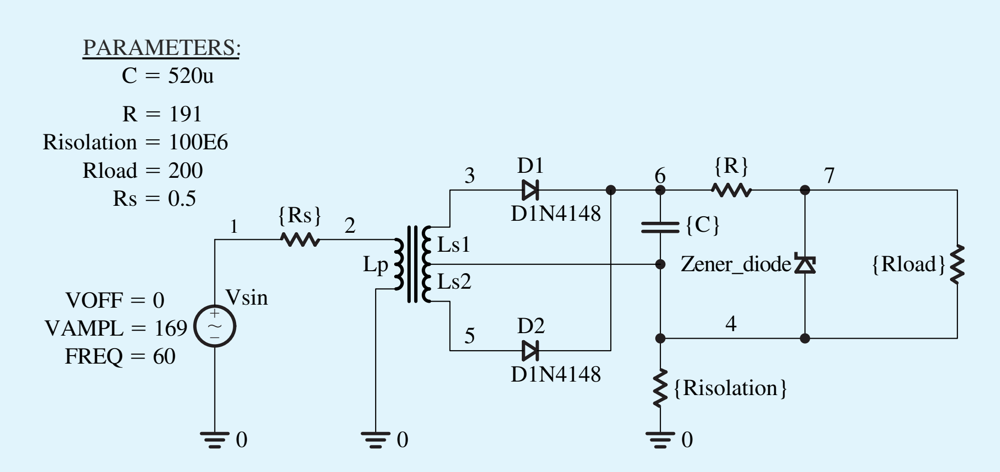


Figure B.13 Schematic capture of the 5-V dc power supply in Example PS.4.1.

An approximate first-cut design can be obtained as follows: The 120-V (rms) supply is stepped down to provide 12-V (peak) sinusoids across each of the secondary windings using a 14:1 turns ratio for the center-tapped transformer. The choice of 12 V is a reasonable compromise between the need to allow for sufficient voltage (above the 5-V output) to operate the rectifier and the regulator, while keeping the PIV ratings of the diodes reasonably low. To determine a value for R, we can use the following expression:

$$R = \frac{V_{C\min} - V_{Z0} - r_z I_{Z\min}}{I_{Z\min} + I_{L\max}}$$

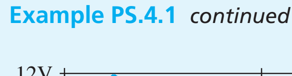
where an estimate for $V_{C\min}$, the minimum voltage across the capacitor, can be obtained by subtracting a diode drop (say, 0.8 V) from 12 V and allowing for a ripple voltage across the capacitor of, say, $V_r = 0.5$ V. Thus, $V_{S\ \min} = 10.7$ V. Furthermore, we note that $I_{L\ \max} = 25$ mA and $I_{Z\min} = 5$ mA, and that $V_Z0 = 4.9$ V and $r_z = 10\ \Omega$. The result is that $R = 191\ \Omega$.

Next, we determine C using a restatement of Eq. (4.33) with V_p/R replaced by the current through the 191-Ω resistor. This current can be estimated by noting that the voltage across C varies from 10.7 V to 11.2 V, and thus has an average value of 10.95 V. Furthermore, the desired voltage across the zener is 5 V. The result is $C = 520\ \mu$F.

Now, with an approximate design in hand, we can proceed with the SPICE simulation. For the zener diode, we use the model of Fig. B.4, and assume (arbitrarily) that D_1 has $I_S = 100$ pA and $n = 0.01$ while D_2 has $I_S = 100$ pA and $n = 1.7$. For the rectifier diodes, we use the commercially available 1N4148 type[2] (with $I_S = 2.682$ nA, $n = 1.836$, $R_S = 0.5664\ \Omega$, $V_0 = 0.5$ V, $C_{j0} = 4$ pF, $m = 0.333$, $\tau_T = 11.54$ ns, $V_{ZK} = 100$ V, $I_{ZK} = 100\ \mu$A).

In PSpice, we perform a transient analysis and plot the waveforms of both the voltage v_C across the smoothing capacitor C and the voltage v_O across the load resistor R_{load}. The simulation results for $R_{\text{load}} = 200\ \Omega$ ($I_{\text{load}} \simeq 25$ mA) are presented in Fig. B.14. Observe that v_C has an average of 10.85 V and a ripple of ± 0.21 V. Thus, $V_1 = 0.42$ V, which is close to the 0.5-V value that we would expect from the chosen value of C. The output voltage v_O is very close to the required 5 V, with v_O varying between 4.957 V and 4.977 V for a ripple of only 20 mV. The variations of v_O with R_{load} are illustrated in Fig. B.15

[2]The 1N4148 model is included in the evaluation (EVAL) library of PSpice.



Example PS.4.1 *continued*

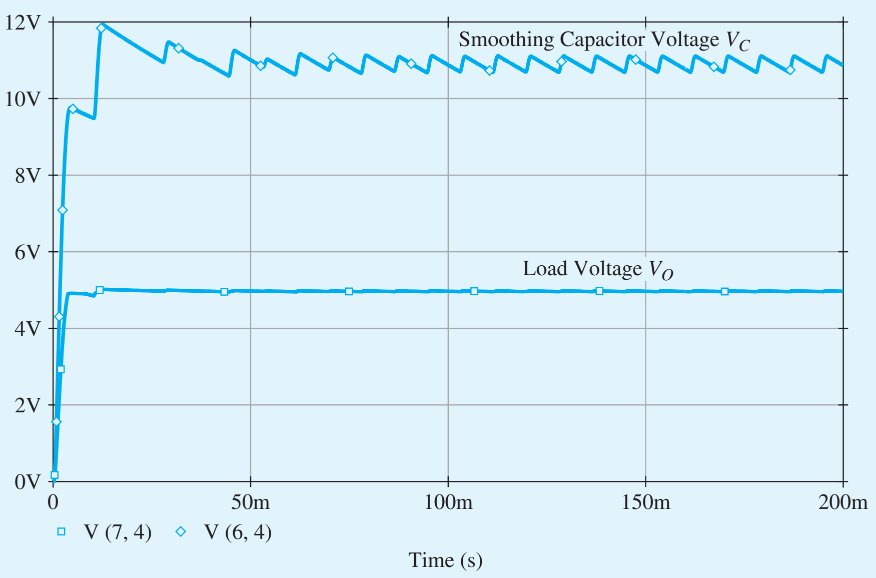


Figure B.14 The voltage v_C across the smoothing capacitor C and the voltage v_O across the load resistor $R_{load} = 200\ \Omega$ in the 5-V power supply of Example PS.4.1.

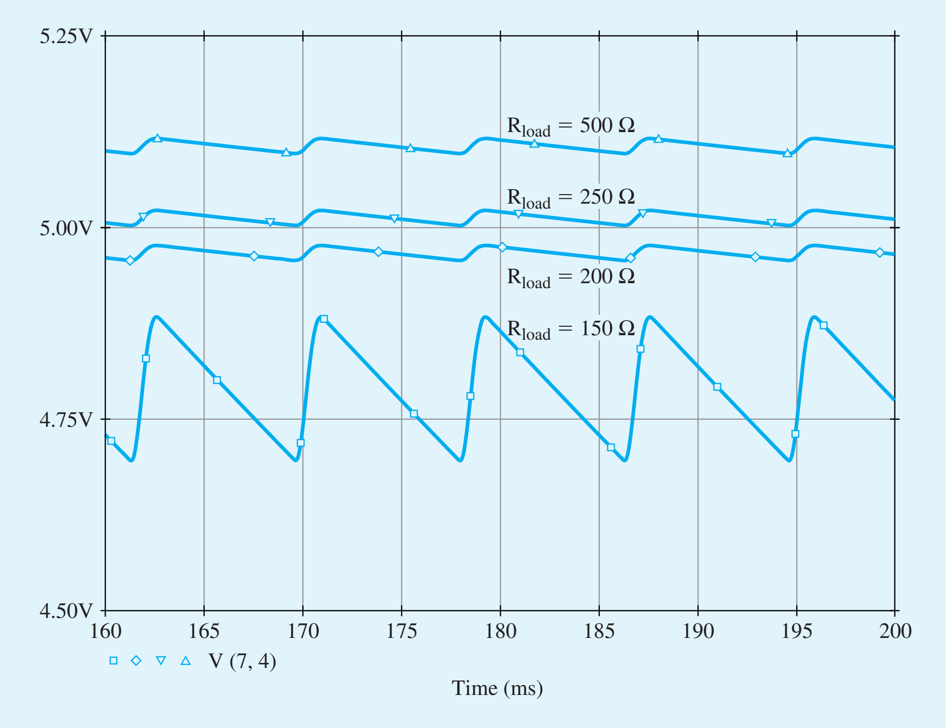


Figure B.15 The output-voltage waveform from the 5-V power supply (in Example PS.4.1) for various load resistances: $R_{load} = 500\ \Omega$, $250\ \Omega$, $200\ \Omega$, and $150\ \Omega$. The voltage regulation is lost at a load resistance of $150\ \Omega$.

for $R_{\text{load}} = 500\,\Omega$, $250\,\Omega$, $200\,\Omega$, and $150\,\Omega$. Accordingly, v_O remains close to the nominal value of 5 V for R_{load} as low as $200\,\Omega(I_{\text{load}} \simeq 25$ mA). For $R_{\text{load}} = 150\,\Omega$ (which implies $I_{\text{load}} \simeq 33.3$ mA, greater than the maximum designed value), we see a significant drop in v_O (to about 4.8 V), as well as a large increase in the ripple voltage at the output (to about 190 mV). This is because the zener regulator is no longer operational; the zener has in fact cut off.

We conclude that the design meets the specifications, and we can stop here. Alternatively, we may consider using further runs of PSpice to help with the task of fine-tuning the design. For instance, we could consider what happens if we use a lower value of C, and so on. We can also investigate other properties of the present design (e.g., the maximum current through each diode) and ascertain whether this maximum is within the rating specified for the diode.

EXERCISE

B.1 Use PSpice to investigate the operation of the voltage doubler whose schematic capture is shown in Fig. B.16(a). Specifically, plot the transient behavior of the voltages v_2 and v_{OUT} when the input is a sinusoid of 10-V peak and 1-kHz frequency. Assume that the diodes are of the 1N4148 type (with $I_S = 2.682$ nA, $n = 1.836$, $R_S = 0.5664\,\Omega$, $V_0 = 0.5$ V, $C_j0 = 4$ pF, $m = 0.333$, $\tau_T = 11.54$ ns, $V_{ZK} = 100$ V, $I_{ZK} = 100\,\mu$A).

Ans. The voltage waveforms are shown in Fig. B.16(b).

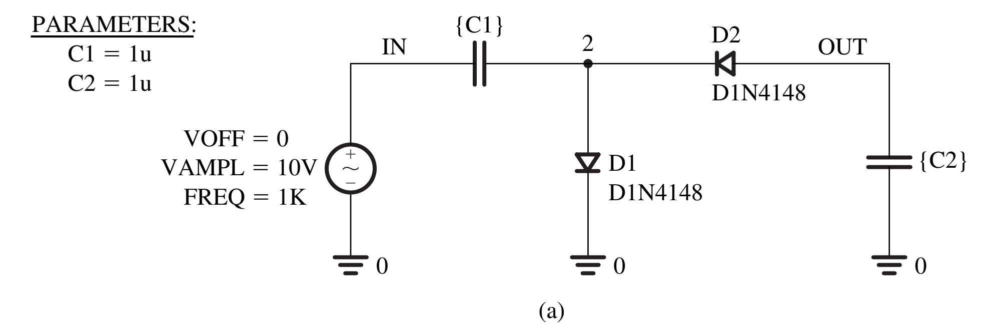
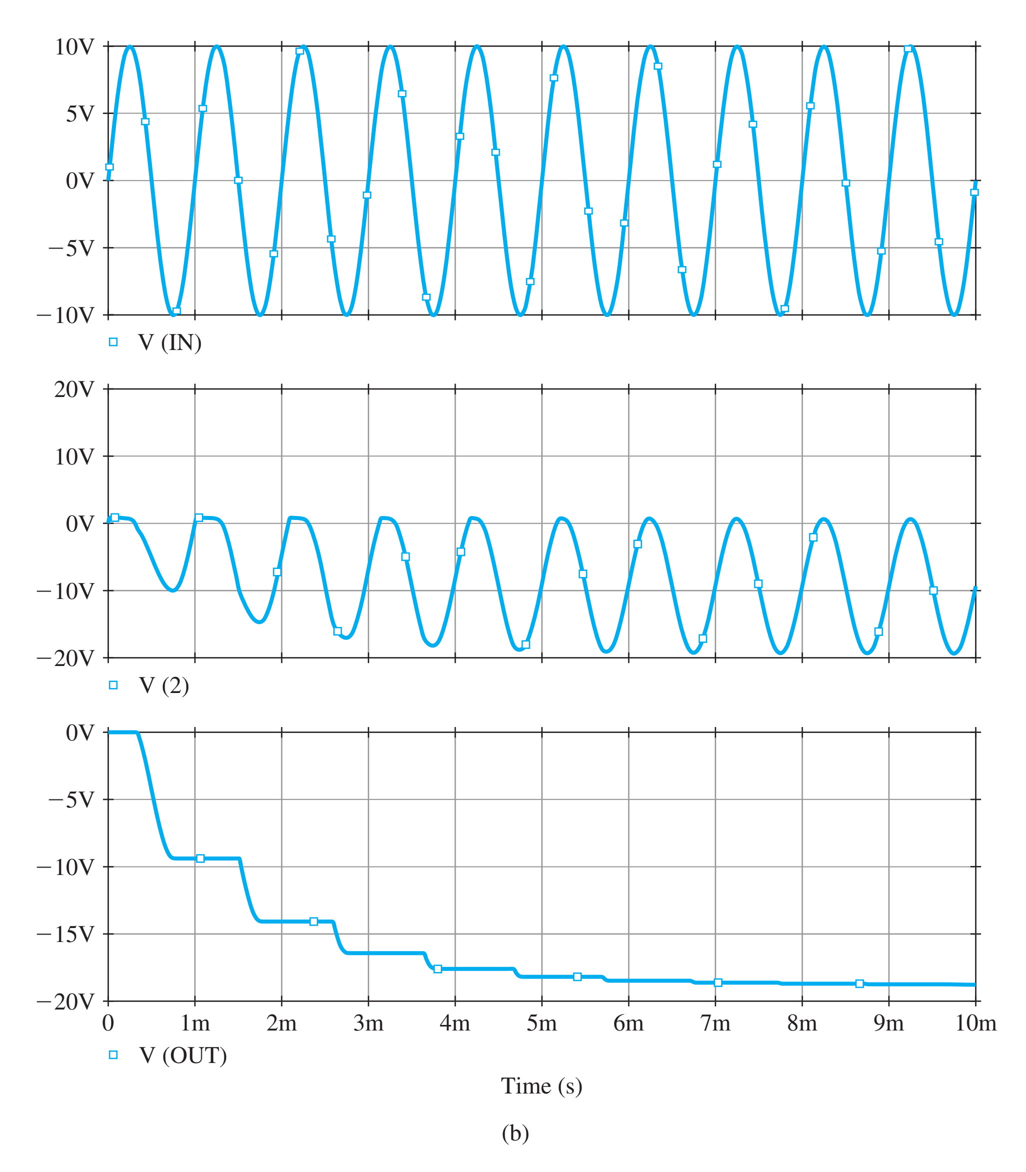


Figure EB.16 **(a)** Schematic capture of the voltage-doubler circuit in Exercise B.1. **(b)** Various voltage waveforms in the voltage-doubler circuit. The top graph displays the input sine-wave voltage signal, the middle graph displays the voltage across diode D_1, and the bottom graph displays the voltage that appears at the output.



(b)

Figure EB.16 *continued*

Example PS.7.1

The CS Amplifier

In this example, we will use PSpice to analyze and verify the design of the CS amplifier whose capture schematic is shown in Fig. B.17.[3] Observe that the MOSFET has its source and body connected in order to cancel the body effect. We will assume a 0.5-μm CMOS technology for the MOSFET and use the SPICE level-1 model parameters listed in Table B.3. We will also assume a signal-source resistance $R_{sig} = 10\ \text{k}\Omega$, a load resistance $R_L = 50\ \text{k}\Omega$, and bypass and coupling capacitors of 10 μF. The targeted specifications for this CS amplifier are a midband gain $A_M = 10$ V/V and a maximum power consumption $P = 1.5$ mW. As should always be the case with computer simulation, we will begin with an approximate pencil-and-paper design. We will then use PSpice to fine-tune our design and to investigate the performance of the final design. In this way, maximum advantage and insight can be obtained from simulation.

With a 3.3-V power supply, the drain current of the MOSFET must be limited to $I_D = P/V_{DD} = 1.5\,\text{mW}/3.3\ \text{V} = 0.45\,\text{mA}$ to meet the power consumption specification. Choosing $V_{OV} = 0.3$ V (a typical value in low-voltage designs) and $V_{DS} = V_{DD}/3$ (to achieve a large signal swing at the output), the MOSFET can now be sized as

$$\frac{W}{L_{\text{eff}}} = \frac{I_D}{\frac{1}{2}k_n' V_{OV}^2\left(1+\lambda V_{DS}\right)} = \frac{0.45\times 10^{-3}}{\frac{1}{2}\left(170.1\times 10^{-6}\right)(0.3)^2[1+0.1(1.1)]} \simeq 53 \qquad \text{(B.19)}$$

where $k_n' = \mu_n C_{ox} = 170.1\ \mu\text{A/V}^2$ (from Table B.3). Here, L_{eff} rather than L is used to more accurately compute I_D. The effect of using W_{eff} rather than W is much less important because typically $W \gg W_{ov}$. Thus, choosing $L = 0.6\ \mu\text{m}$ results in $L_{\text{eff}} = L - 2L_{ov} = 0.44\ \mu\text{m}$ and $W = 23.3\ \mu\text{m}$. Note that we chose L slightly larger than $L_{\min}$. This is a common practice in the design of analog ICs to minimize the effects of fabrication nonidealities on the actual value of L. As shown in the text, this is particularly important when

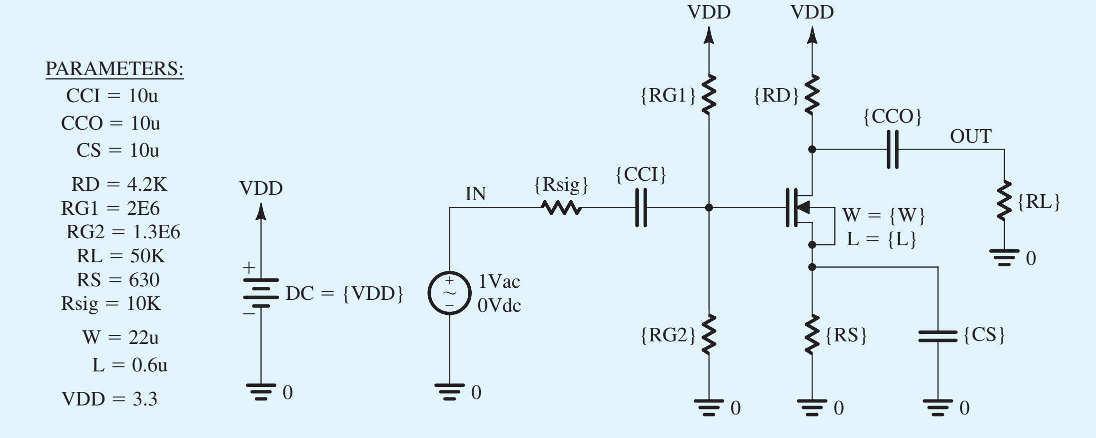


Figure B.17 Schematic capture of the CS amplifier in Example PS.7.1.

[3]The reader is reminded that the schematic capture diagrams and the corresponding PSpice simulation files of all SPICE examples in this book can be found on the text's website (www.oup.com/us/sedrasmith). In these schematics (as shown in Fig. B.17), we used variable parameters to enter the values of the various circuit components, including the dimensions of the MOSFET. This will allow the reader to investigate the effect of changing component values by simply changing the corresponding parameter values.

Example PS.7.1 *continued*

the circuit performance depends on the matching between the dimensions of two or more MOSFETs (e.g., in the current-mirror circuits studied in Chapter 8).

Next, R_D is calculated based on the desired voltage gain:

$$|A_v| = g_m(R_D||R_L||r_o) = 10\,\text{V/V} \Rightarrow R_D \simeq 4.2\,\text{k}\Omega \tag{B.20}$$

where $g_m = 3.0$ mA/V and $r_o = 22.2$ kΩ. Hence, the output bias voltage is $V_O = V_{DD} - I_D R_D = 1.39$ V. An $R_S = (V_O - V_{DD}/3)/I_D = 630\,\Omega$ is needed to bias the MOSFET at a $V_{DS} = V_{DD}/3$. Finally, resistors $R_{G1} = 2\,\text{M}\Omega$ and $R_{G2} = 1.3\,\text{M}\Omega$ are chosen to set the gate bias voltage at $V_G = I_D R_S + V_{OV} + V_{tn} \simeq 1.29$ V. Using large values for these gate resistors ensures that both their power consumption and the loading effect on the input signal source are negligible. Note that we neglected the body effect in the expression for V_G to simplify our hand calculations.

We will now use PSpice to verify our design and investigate the performance of the CS amplifier. We begin by performing a bias-point simulation to verify that the MOSFET is properly biased in the saturation region and that the dc voltages and currents are within the desired specifications. Based on this simulation, we have decreased the value of W to 22 μm to limit I_D to about 0.45 mA. Next, to measure the midband gain A_M and the 3-dB frequencies[4] f_L and f_H, we apply a 1-V ac voltage at the input, perform an ac-analysis simulation, and plot the output-voltage magnitude (in dB) versus frequency as shown in Fig. B.18. This corresponds to the magnitude response of the CS amplifier because we chose a 1-V input signal.[5] Accordingly, the midband gain is $A_M = 9.55$ V/V and the 3-dB bandwidth is $BW = f_H - f_L \simeq 122.1$ MHz. Figure B.18 further shows that the gain begins to fall off at about 300 Hz but flattens out again at about 10 Hz. This flattening in the gain at low frequencies is due to a real transmission zero[6] introduced in the transfer function of the amplifier by R_S together with C_S. This zero occurs at a frequency $f_Z = 1/(2\pi R_S C_S) = 25.3$ Hz, which is typically between the break frequencies f_{P2} and f_{P3} derived in Section 10.1.1. So, let us now verify this phenomenon by resimulating the CS amplifier with a $C_S = 0$ (i.e., removing C_S) in order to move f_Z to infinity and remove its effect. The corresponding frequency response is plotted also in Fig. B.18. As expected, with $C_S = 0$, we do not observe any flattening in the low-frequency response of the amplifier. However, because the CS amplifier now includes a source resistor R_S, A_M has dropped by a factor of 2.6. This factor is approximately equal to $(1 + g_m R_S)$, as expected from our study of the CS amplifier with a source-degeneration resistance in Section 7.3.4. Note that the bandwidth BW has increased by approximately the same factor as the drop in gain A_M. As we will learn in Chapter 11 when we study negative feedback, the source-degeneration resistor R_S provides negative feedback, which allows us to trade off gain for wider bandwidth.

To conclude this example, we will demonstrate the improved bias stability achieved when a source resistor R_S is used (see the discussion in Section 7.1.1 & 7.1.7). Specifically, we will change (in the MOSFET level-1 model for part NMOS0P5) the value of the zero-bias threshold voltage parameter VTO by ±15% and perform a bias-point simulation in PSpice. Table B.5 shows the corresponding variations in I_D and V_O for the case in which $R_S = 630\,\Omega$. For the case without source degeneration, we use an $R_S = 0$ in the

---

[4]No detailed knowledge of frequency-response calculations is required for this example; all that is needed is Section 7.5.3. Nevertheless, after the study of the frequency response of the CS amplifier in Sections 10.1 through 10.3, the reader will benefit by returning to this example and using PSpice to experiment further with the circuit.
[5]The reader should not be alarmed about the use of such a large signal amplitude. Recall that in a small-signal (ac) simulation, SPICE first finds the small-signal equivalent circuit at the bias point and then analyzes this linear circuit. Such ac analysis can, of course, be done with any ac signal amplitude. However, a 1-V ac input is convenient to use because the resulting ac output corresponds to the voltage gain of the circuit.
[6]Readers who have not yet studied poles and zeros can skip these few sentences.

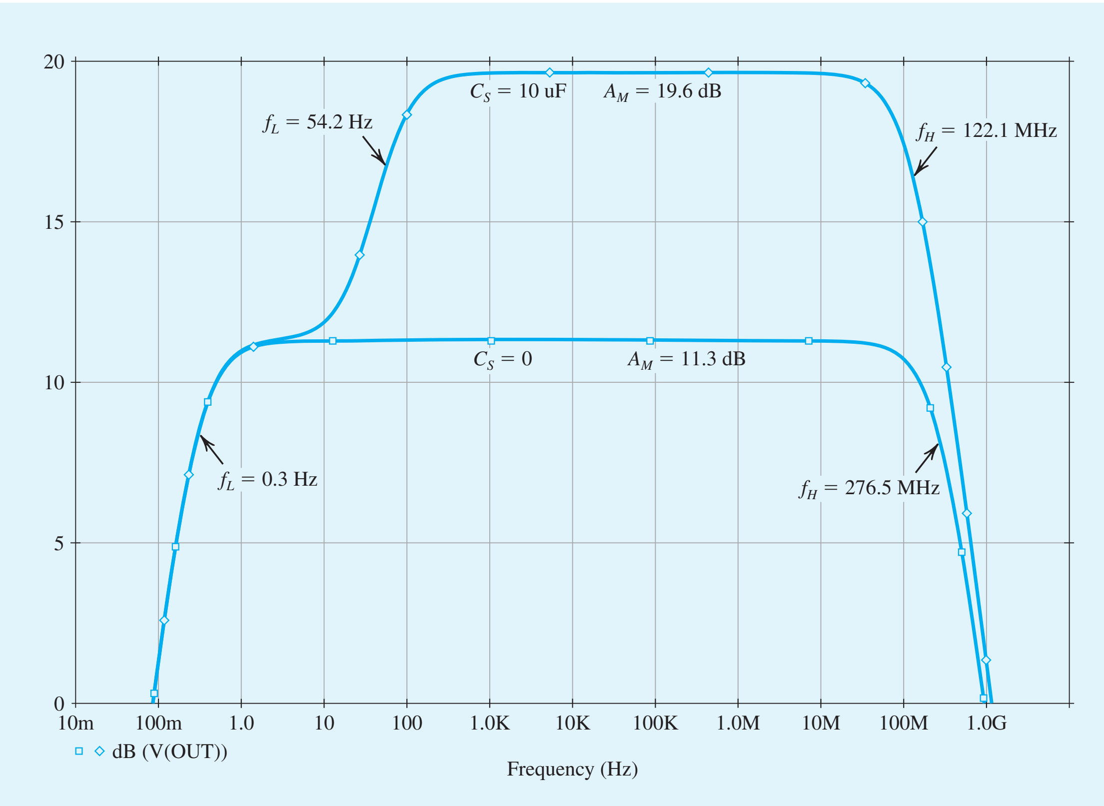


Figure B.18 Frequency response of the CS amplifier in Example PS.7.1 with $C_S = 10\,\mu\text{F}$ and $C_S = 0$ (i.e., C_S removed).

schematic of Fig. B.17. Furthermore, to obtain the same I_D and V_O in both cases (for the nominal threshold voltage $V_{t0} = 0.7$ V), we use an $R_{G2} = 0.88\ \text{M}\Omega$ to reduce V_G to around $V_{OV} + V_{tn} = 1$ V. The corresponding variations in the bias point are shown in Table B.5. Accordingly, we see that the source-degeneration resistor makes the bias point of the CS amplifier less sensitive to changes in the threshold voltage. In fact, the reader can show for the values displayed in Table B.5 that the variation in bias current ($\Delta I/I$) is reduced by approximately the same factor, $(1 + g_m R_S)$. However, unless a large bypass capacitor C_S is used, this reduced sensitivity comes at the expense of a reduction in the midband gain (as we observed in this example when we simulated the frequency response of the CS amplifier with a $C_S = 0$).

Table B.5 Variations in the Bias Point with the MOSFET Threshold Voltage

	$R_S = 630\ \Omega$		$R_S = 0$	
V_{tn0}	I_D (mA)	V_O (V)	I_D (mA)	V_O (V)
0.60	0.56	0.962	0.71	0.33
0.7	0.46	1.39	0.45	1.40
0.81	0.36	1.81	0.21	2.40

Example PS.6.1

Dependence of the BJT β on the Bias Current

In this example, we use PSpice to simulate the dependence of β_{dc} on the collector bias current for the Q2N3904 discrete BJT (from Fairchild Semiconductor) whose model parameters are listed in Table B.6 and are available in PSpice.[7] As shown in the schematic capture[8] of Fig. B.19, the V_{CE} of the BJT is fixed using a constant voltage source (in this example, $V_{CE} = 2$ V) and a dc current source I_B is applied at the base. To illustrate the dependence of β_{dc} on the collector current I_C, we perform a dc-analysis simulation in which the sweep variable is the current source I_B. The β_{dc} of the BJT, which corresponds to the ratio of the collector current I_C to the base current I_B, can then be plotted versus I_C using Probe (the graphical interface of PSpice), as shown in Fig. B.20. We see that to operate at the maximum value of β_{dc} (i.e., $\beta_{dc} = 163$), at $V_{CE} = 2$ V, the BJT must be biased at an $I_C = 10$ mA. Since increasing the bias current of a transistor increases the power dissipation, it is clear from Fig. B.20 that the choice of current I_C is a trade-off between the current gain β_{dc} and the power dissipation. Generally speaking, the optimum I_C depends on the application and technology in hand. For example, for the Q2N3904 BJT operating at $V_{CE} = 2$ V, decreasing I_C by a factor of 20 (from 10 mA to 0.5 mA) results in a drop in β_{dc} of about 25% (from 163 to 123)

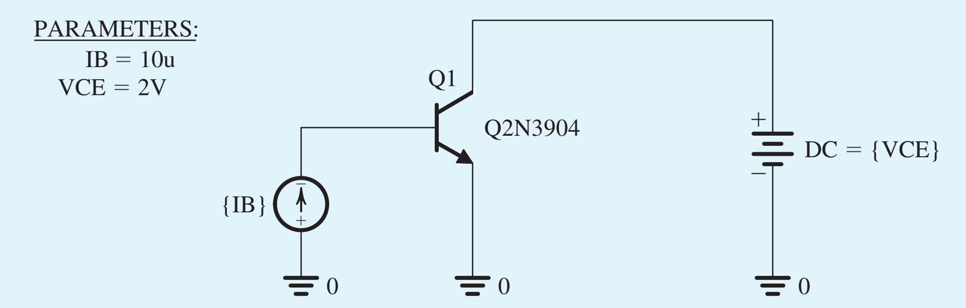


Figure B.19 The PSpice test bench used to demonstrate the dependence of β_{dc} on the collector bias current I_C for the Q2N3904 discrete BJT (Example PS.6.1).

Table B.6 Spice Model Parameters of the Q2N3904 Discrete BJT

IS = 6.734F	XTI = 3	EG = 1.11	VAF = 74.03	BF = 416.4	NE = 1.259	ISE = 6.734F
IKF = 66.78M	XTB = 1.5	BR = .7371	NC = 2	ISC = 0	IKR = 0	RC = 1
CJC = 3.638P	MJC = .3085	VJC = .75	FC = .5	CJE = 4.493P	MJE = .2593	VJE = .75
TR = 239.5N	TF = 301.2P	ITF = .4	VTF = 4	XTF = 2	RB = 10	

[7]The Q2N3904 model is included in the evaluation (EVAL) library of PSpice, which is available on the website accompanying this book.

[8]The reader is reminded that the schematic diagrams and the corresponding PSpice simulation files of all SPICE examples in this book can be found on the text's website (www.oup.com/us/sedrasmith). In these schematics (as shown in Fig. B.19), we use variable parameters to enter the values of the various circuit components. This allows one to investigate the effect of changing component values by simply changing the corresponding parameter values.

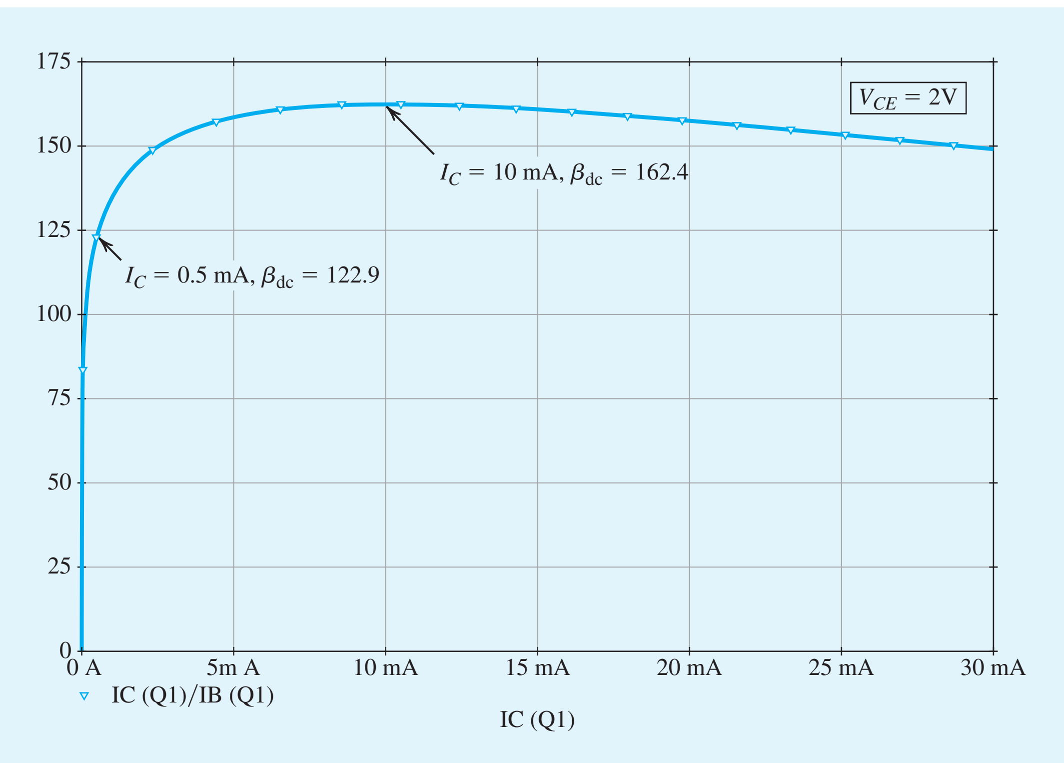


Figure B.20 Dependence of β_{dc} on I_C (at $V_{CE} = 2$ V) in the Q2N3904 discrete BJT (Example PS.6.1).

Example PS.7.2

The CE Amplifier with Emitter Resistance

In this example, we use PSpice to analyze and verify the design of the CE amplifier. A schematic capture of the CE amplifier is shown in Fig. B.21. We will use part Q2N3904 for the BJT and a ±5-V power supply. We will also assume a signal source resistor $R_{sig} = 10$ kΩ, a load resistor $R_L = 10$ kΩ, and bypass and coupling capacitors of 10 μF. To enable us to investigate the effect of including a resistance in the signal path of the emitter, a resistor R_{ce} is connected in series with the emitter bypass capacitor C_E. Note that the roles of R_E and R_{ce} are different. Resistor R_E is the **dc emitter-degeneration resistor** because it appears in the dc path between the emitter and ground. It is therefore used to help stabilize the bias point for the amplifier. The equivalent resistance $R_e = R_E \| R_{ce}$ is the **small-signal emitter-degeneration resistance** because it appears in the ac (small-signal) path between the emitter and ground and helps stabilize the gain of the amplifier. In this example, we will investigate the effects of both R_E and R_e on the performance of the CE amplifier. However, as should always be the case with computer simulation, we will begin with an approximate pencil-and-paper design. In this way, maximum advantage and insight can be obtained from simulation.

Based on the plot of β_{dc} versus I_C in Fig. B.20, a collector bias current I_C of 0.5 mA is selected for the BJT, resulting in $\beta_{dc} = 123$. This choice of I_C is a reasonable compromise between power dissipation and current gain. Furthermore, a collector bias voltage V_C of 0 V (i.e., at the mid–supply rail) is selected to

Example PS.7.2 *continued*

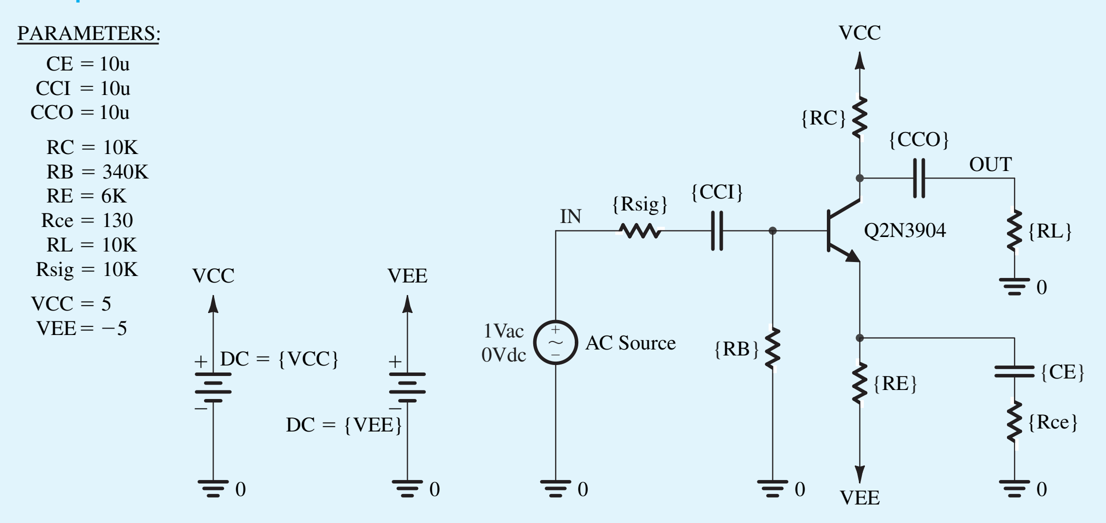


Figure B.21 Schematic capture of the CE amplifier in Example PS.7.2.

achieve a high signal swing at the amplifier output. For $V_{CE} = 2$ V, the result is that $V_E = -2$ V requires bias resistors with values

$$R_C = \frac{V_{CC} - V_C}{I_C} = 10\,\text{k}\Omega$$

and

$$R_E = \frac{V_E - V_{EE}}{I_C} = 6\,\text{k}\Omega$$

Assuming $V_{BE} = 0.7$ V and using $\beta_{\text{dc}} = 123$, we can determine

$$R_B = -\frac{V_B}{I_B} = -\frac{0 - (V_{BE} + V_E)}{I_C/\beta_{\text{dc}}} = 320\,\text{k}\Omega$$

Next, the formulas of Section 7.3.4 can be used to determine the input resistance R_{in} and the midband voltage gain $|A_M|$ of the CE amplifier:

$$R_{\text{in}} = R_B \,\|\, (\beta_{\text{ac}} + 1) \,\|\, (r_e + R_e) \tag{B.21}$$

$$|A_M| = \left| -\frac{R_{\text{in}}}{R_{\text{sig}} + R_{\text{in}}} \times \frac{R_C \,\|\, R_L}{r_e + R_e} \right| \tag{B.22}$$

For simplicity, we will assume $\beta_{\text{ac}} \simeq \beta_{\text{dc}} = 123$, resulting in

$$r_e = \left(\frac{\beta_{\text{ac}}}{\beta_{\text{ac}} + 1}\right)\left(\frac{V_T}{I_C}\right) = 49.6\,\Omega$$

Thus, with no small-signal emitter degeneration (i.e., $R_{ce} = 0$), $R_{\text{in}} = 6.1\,\text{k}\Omega$ and $|A_M| = 38.2$ V/V. Using Eq. (B.22) and assuming R_B is large enough to have a negligible effect on R_{in}, it can be shown that

the emitter-degeneration resistor R_e decreases the voltage gain $|A_M|$ by a factor of

$$\frac{1+\dfrac{R_e}{r_e}+\dfrac{R_{\text{sig}}}{r_\pi}}{1+\dfrac{R_{\text{sig}}}{r_\pi}}$$

Therefore, to limit the reduction in voltage gain to a factor of 2, we will select

$$R_e = r_e + \frac{R_{\text{sig}}}{\beta_{\text{ac}}+1} \tag{B.23}$$

Thus, $R_{ce} \simeq R_e = 130\,\Omega$. Substituting this value in Eqs. (B.21) and (B.22) shows that R_{in} increases from $6.1\,\text{k}\Omega$ to $20.9\,\text{k}\Omega$ while $|A_M|$ drops from 38.2 V/V to 18.8 V/V.

We will now use PSpice to verify our design and investigate the performance of the CE amplifier. We begin by performing a bias-point simulation to verify that the BJT is properly biased in the active region and that the dc voltages and currents are within the desired specifications. Based on this simulation, we have increased the value of R_B to $340\,\text{k}\Omega$ in order to limit I_C to about 0.5 mA while using a standard 1% resistor value (Appendix H). Next, to measure the midband gain A_M and the 3-dB frequencies[9] f_L and f_H, we apply a 1-V ac voltage at the input, perform an ac-analysis simulation, and plot the output-voltage magnitude (in dB) versus frequency as shown in Fig. B.22. This corresponds to the magnitude response of the CE amplifier because we chose a 1-V input signal.[10] Accordingly, with no emitter degeneration, the midband gain is $|A_M| = 38.5$ V/V $= 31.7$ dB and the 3-dB bandwidth is $BW = f_H - f_L = 145.7\,\text{kHz}$. Using an R_{ce} of $130\,\Omega$ results in a drop in the midband gain $|A_M|$ by a factor of 2 (i.e., 6 dB). Interestingly, however, BW has now increased by approximately the same factor as the drop in $|A_M|$. As we learned in Chapter 10 in our study of negative feedback, the emitter-degeneration resistor R_{ce} provides negative feedback, which allows us to trade off gain for other desirable properties, such as a larger input resistance and a wider bandwidth.

To conclude this example, we will demonstrate the improved bias-point (or dc operating-point) stability achieved when an emitter resistor R_E is used (see the discussion in Section 7.4.2). Specifically, we will increase/decrease the value of the parameter BF (i.e., the ideal maximum forward current gain) in the SPICE model for part Q2N3904 by a factor of 2 and perform a bias-point simulation. The corresponding change in BJT parameters (β_{dc} and β_{ac}) and bias-point (including I_C and CE) are presented in Table B.7 for the case of $R_E = 6\,\text{k}\Omega$. Note that β_{ac} is not equal to β_{dc} as we assumed, but is slightly larger. For the case without emitter degeneration, we will use $R_E = 0$ in the schematic of Fig. B.21. Furthermore, to maintain the same I_C and V_C in both cases at the values obtained for nominal BF, we use $R_B = 1.12\ \text{M}\Omega$ to limit I_C to approximately 0.5 mA. The corresponding variations in the BJT bias point are also shown in Table B.7. Accordingly, we see that emitter degeneration makes the bias point of the CE amplifier much less sensitive to changes in β. However, unless a large bypass capacitor C_E is used, this reduced bias sensitivity comes at the expense of a reduction in the midband gain (as we observed in this example when we simulated the frequency response of the CE amplifier with an $R_e = 130\,\Omega$).

[9]No detailed knowledge of frequency-response calculations is required for this example; all that is needed is Section 7.4.2. Nevertheless, after the study of the frequency of the CE amplifier in Sections 10.1 through 10.3, the reader will benefit by returning to this example to experiment further with the circuit using PSpice.

[10]The reader should not be alarmed about the use of such a large signal amplitude. Recall that in a small-signal (ac) simulation, SPICE first finds the small-signal equivalent circuit at the dc bias point and then analyzes this linear circuit. Such ac analysis can, of course, be done with any ac signal amplitude. However, a 1-V ac input is convenient to use because the resulting ac output corresponds to the voltage gain of the circuit.

Example PS.7.2 *continued*

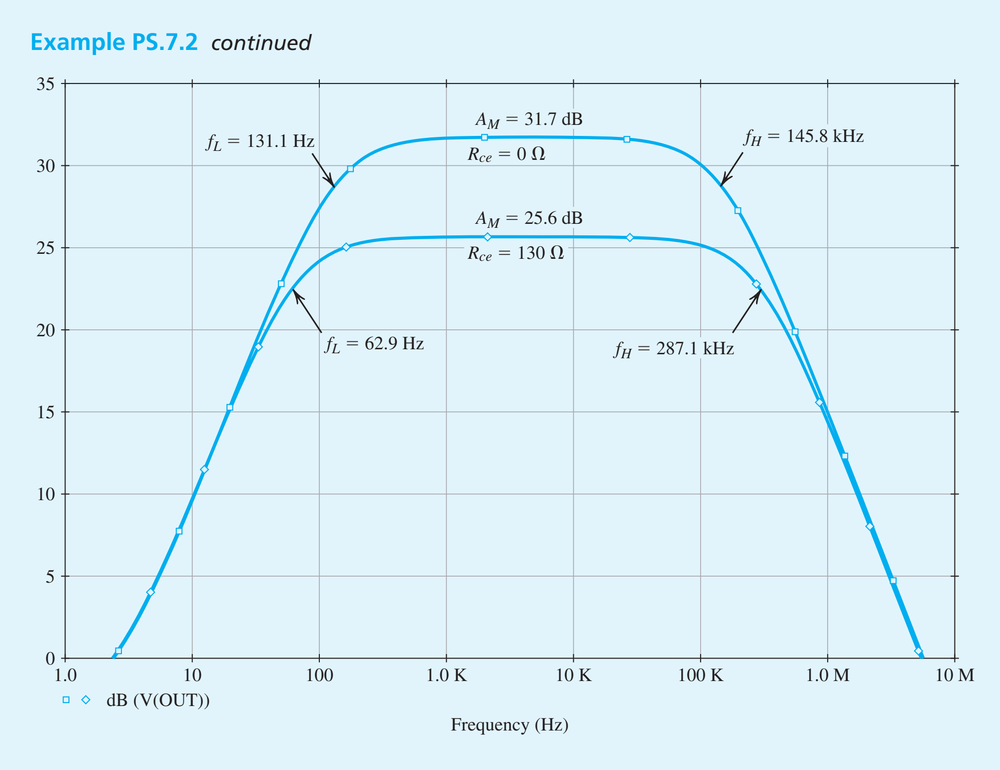


Figure B.22 Frequency response of the CE amplifier in Example PS.7.3 with $R_{ce} = 0$ and $R_{ce} = 130\ \Omega$.

Table B.7 Variations in the Bias Point of the CE Amplifier with the SPICE Model-Parameter BF of BJT

	$R_E = 6$ k				$R_E = 0$			
BF (in SPICE)	β_{ac}	β_{dc}	I_C (mA)	V_C (V)	β_{ac}	β_{dc}	I_C (mA)	V_C (V)
208	106	94.9	0.452	0.484	109	96.9	0.377	1.227
416.4 (nominal value)	143	123	0.494	0.062	148	127	0.494	0.060
832	173	144	0.518	−0.183	181	151	0.588	−0.878

Example PS.8.1

The CMOS CS Amplifier

In this example, we will use PSpice to compute the dc transfer characteristic of the CS amplifier whose capture schematic is shown in Fig. B.23. We will assume a 5-µm CMOS technology for the MOSFETs and use parts NMOS5P0 and PMOS5P0 whose SPICE level-1 parameters are listed in Table B.3. To specify the dimensions of the MOSFETs in PSpice, we will use the multiplicative factor m together with the channel length L and the channel width W. The MOSFET parameter m, whose default value is 1, is used in SPICE to specify the number of MOSFETs connected in parallel. As depicted in Fig. B.24, a wide transistor with channel length L and channel width $m \times W$ can be implemented using m narrower transistors

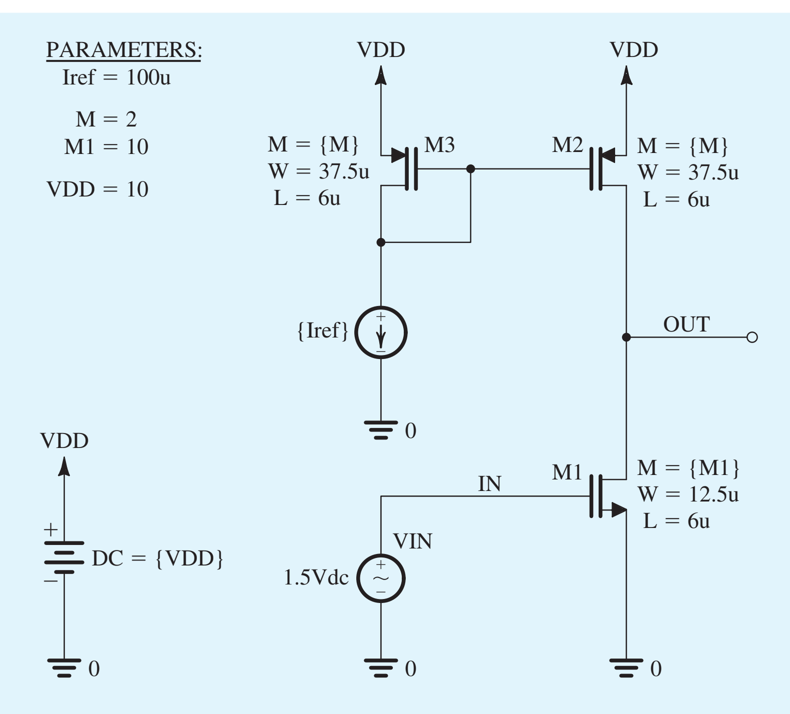


Figure B.23 Schematic capture of the CS amplifier in Example PS.8.1.

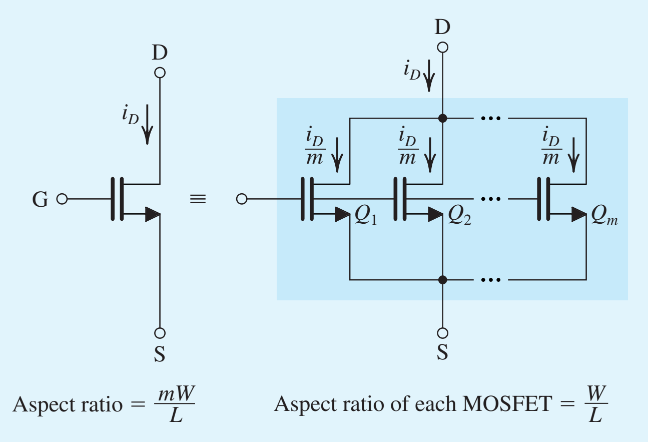


Figure B.24 Transistor equivalency.

in parallel, each having a channel length L and a channel width W. Thus, neglecting the channel-length modulation effect, the drain current of a MOSFET operating in the saturation region can be expressed as

$$I_D = \frac{1}{2}\mu C_{ox} m \frac{W}{L_{\text{eff}}} V_{OV}^2 \tag{B.24}$$

where L_{eff} rather than L is used to more accurately estimate the drain current.

The CS amplifier in Fig. B.23 is designed for a bias current of 100 μA assuming a reference current $I_{\text{ref}} = 100\,\mu\text{A}$ and $V_{DD} = 10$ V. The current mirror transistors M_2 and M_3 are sized for $V_{OV2} = V_{OV3} = 1$ V,

Example PS.8.1 *continued*

while the input transistor M_1 is sized for $V_{OV1} = 0.5$ V. Note that a smaller overdrive voltage is selected for M_1 to achieve a larger voltage gain G_v for the CS amplifier, since

$$G_v = -g_{m1}R'_L = -g_{m1}(r_{o1} \| r_{o2}) = -\frac{2}{V_{OV1}}\left(\frac{V_{An}V_{Ap}}{V_{An}+V_{Ap}}\right) \tag{B.25}$$

where V_{An} and V_{Ap} are the magnitudes of the Early voltages of, respectively, the NMOS and PMOS transistors. Unit-size transistors are used with $W/L = 12.5$ μm/6 μm for the NMOS devices and $W/L = 37.5$ μm/6 μm for the PMOS devices. Thus, using Eq. (B.24) together with the 5-μm CMOS process parameters in Table B.4, we find $m_1 = 10$ and $m_2 = m_3 = 2$ (rounded to the nearest integer). Furthermore, Eq. (B.25) gives $G_v = -100$ V/V.

To compute the dc transfer characteristic of the CS amplifier, we perform a dc analysis in PSpice with V_{IN} swept over the range 0 to V_{DD} and plot the corresponding output voltage V_{OUT}. Figure B.25(a) shows the resulting transfer characteristic. The slope of this characteristic (i.e., dV_{OUT}/dV_{IN}) corresponds to the gain of the amplifier. The high-gain segment is clearly visible for V_{IN} around 1.5 V. This corresponds to an overdrive voltage for M_1 of $V_{OV1} = V_{IN} - V_{tn} = 0.5$ V, as desired. To examine the high-gain region more closely, we repeat the dc sweep for V_{IN} between 1.3 V and 1.7 V. The resulting transfer characteristic is plotted in Fig. B.25 (b, middle curve). Using the Probe graphical interface of PSpice, we find that the linear region of this dc transfer characteristic is bounded approximately by V_{IN} = 1.465 V and V_{IN} = 1.539 V. The corresponding values of V_{OUT} are 8.838 V and 0.573 V. These results are close to the expected values. Specifically, transistors M_1 and M_2 will remain in the saturation region and, hence, the amplifier will operate in its linear region if $V_{OV1} \le V_{OUT} \le V_{DD} - V_{OV2}$ or $0.5 \text{ V} \le V_{OUT} \le 9$ V. From the results above, the voltage gain G_v (i.e., the slope of the linear segment of the dc transfer characteristic) is approximately −112 V/V, which is reasonably close to the value obtained by hand analysis.

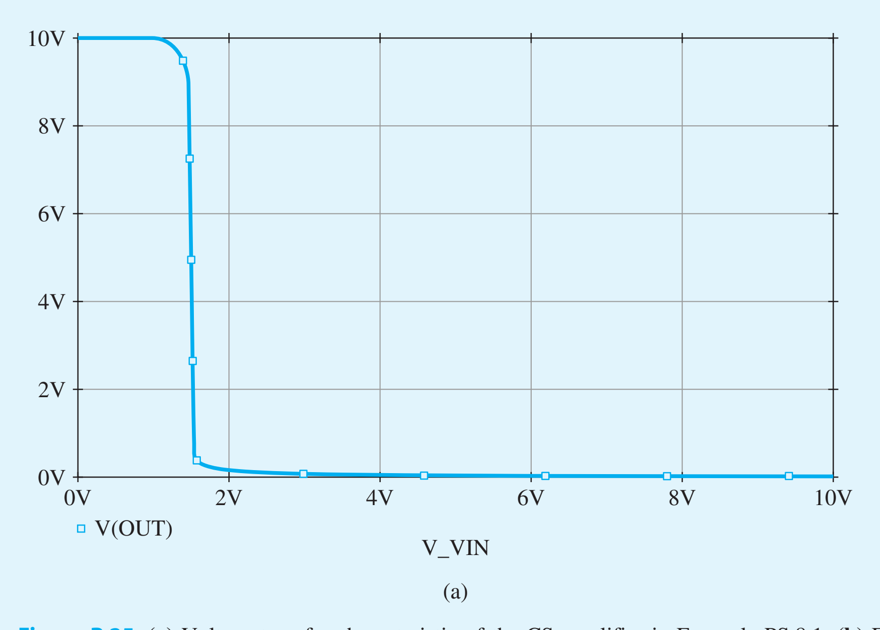


Figure B.25 **(a)** Voltage transfer characteristic of the CS amplifier in Example PS.8.1. **(b)** Expanded view of the transfer characteristic in the high-gain region. Also shown are the transfer characteristics where process variations cause the width of transistor M_1 to change by +15% and –15% from its nominal value of $W_1 = 12.5$ μm.

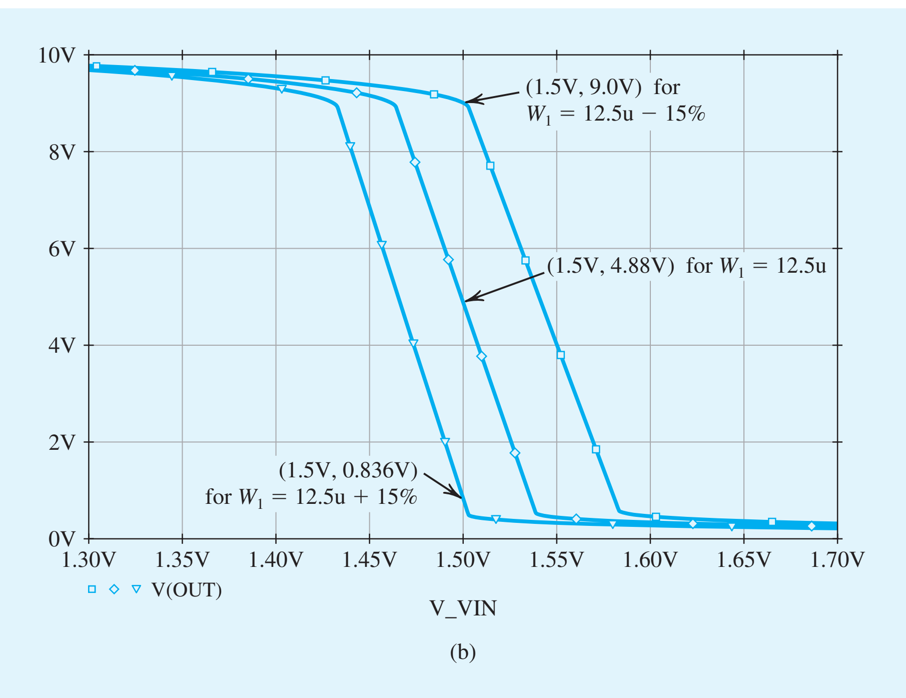


Figure B.25 *continued*

Note from the dc transfer characteristic in Fig. B.25(b) that for an input dc bias of $V_{IN} = 1.5$ V, the output dc bias is $V_{OUT} = 4.88$ V. This choice of V_{IN} maximizes the available signal swing at the output by setting V_{OUT} at the middle of the linear segment of the dc transfer characteristic. However, because of the high resistance at the output node (or, equivalently, because of the high voltage gain), this value of V_{OUT} is highly sensitive to the effect of process and temperature variations on the characteristics of the transistors. To illustrate this point, consider what happens when the width of M_1 (i.e., W_1, which is normally 12.5 μm) changes by ±15%. The corresponding dc transfer characteristics are shown in Fig. B.25(b). Accordingly, when $V_{IN} = 1.5$ V, V_{OUT} will drop to 0.84 V if W_1 increases by 15% and will rise to 9.0 V if W_1 decreases by 15%. In practical circuit implementations, this problem is circumvented by using negative feedback to accurately set the dc bias voltage at the output of the amplifier and, hence, to reduce the sensitivity of the circuit to process variations. We studied negative feedback in Chapter 11.

Example PS.9.1

A Multistage Differential BJT Amplifier

The schematic capture of the multistage op-amp circuit analyzed in Examples 8.1 and 8.7 is shown in Fig. B.26.[11] Observe the manner in which the differential signal input V_d and the common-mode input voltage V_{CM} are applied. Such an input bias configuration for an op-amp circuit was presented and used

[11]This circuit cannot be simulated using the student evaluation version of PSpice that is included on the website accompanying this book. This is because, in this free version of PSpice, circuit simulation is restricted to circuits with no more than 10 transistors.

Example PS.9.1 *continued*

in Example PS.2.2. In the following simulations, we will use parts Q2N3904 and Q2N3906 (from Fairchild Semiconductor) for the *npn* and *pnp* BJTs, respectively. The model parameters of these discrete BJTs are listed in Table B.8 and are available in PSpice.

Table B.8 Spice Model Parameters of the Q2N3904 and Q2N3906 Discrete BJTs

Q2N3904 Discrete BJT						
IS = 6.734f	XTI = 3	EG = 1.11	VAF = 74.03	BF = 416.4	NE = 1.259	ISE = 6.734f
IKF = 66.78m	XTB = 1.5	BR = .7371	NC = 2	ISC = 0	IKR = 0	RC = 1
CJC = 3.638p	MJC = .3085	VJC = .75	FC = .5	CJE = 4.493p	MJE = .2593	VJE = .75
TR = 239.5n	TF = 301.2p	ITF = .4	VTF = 4	XTF = 2	RB = 10	
Q2N3906 Discrete BJT						
IS = 1.41f	XTI = 3	EG = 1.11	VAF = 18.7	BF = 180.7	NE = 1.5	ISE = 0
IKF = 80m	XTB = 1.5	BR = 4.977	NC = 2	ISC = 0	IKR = 0	RC = 2.5
CJC = 9.728p	MJC = .5776	VJC = .75	FC = .5	CJE = 8.063p	MJE = .3677	VJE = .75
TR = 33.42n	TF = 179.3p	ITF = .4	VTF = 4	XTF = 6	RB = 10	

Table B.9 DC Collector Currents of the Op-Amp Circuit in Fig. B.26 as Computed by Hand Analysis (Example 8.6) and by PSpice

	Collector Currents (mA)		
Transistor	Hand Analysis (Example 8.6)	PSpice	Error (%)
Q_1	0.25	0.281	–11.0
Q_2	0.25	0.281	–11.0
Q_3	0.5	0.567	–11.8
Q_4	1.0	1.27	–21.3
Q_5	1.0	1.21	–17.4
Q_6	2.0	2.50	–20.0
Q_7	1.0	1.27	–21.3
Q_8	5.0	6.17	–18.9
Q_9	0.5	0.48	+4.2

In PSpice, the common-mode input voltage V_{CM} of the op-amp circuit is set to 0 V (i.e., to the average of the dc power-supply voltages V_{CC} and V_{EE}) to maximize the available input signal swing. A bias-point simulation is performed to determine the dc operating point. Table B.9 summarizes the value of the dc collector currents as computed by PSpice and as calculated by the hand analysis in Example 8.6. Recall that our hand analysis assumed both β and the Early voltage V_A of the BJTs to be infinite. However, our SPICE simulations in Example PS.6.1 (where we investigated the dependence of β on the collector current I_C) indicate that the Q2N3904 has $\beta \simeq 125$ at $I_C = 0.25$ mA. Furthermore, its forward Early voltage (SPICE parameter VAF) is 74 V, as given in Table B.8. Nevertheless, we observe from Table B.9 that the largest error in the calculation of the dc bias currents is on the order of 20%. Accordingly, we can conclude that a quick hand analysis using gross approximations can still yield reasonable results for a preliminary estimate and, of course, hand analysis yields much insight into the circuit operation. In addition to the dc bias currents listed in Table B.9, the bias-point simulation in PSpice shows that the output dc offset (i.e., V_{OUT} when $V_d = 0$) is 3.62 V and that the input bias current I_{B1} is 2.88 μA.

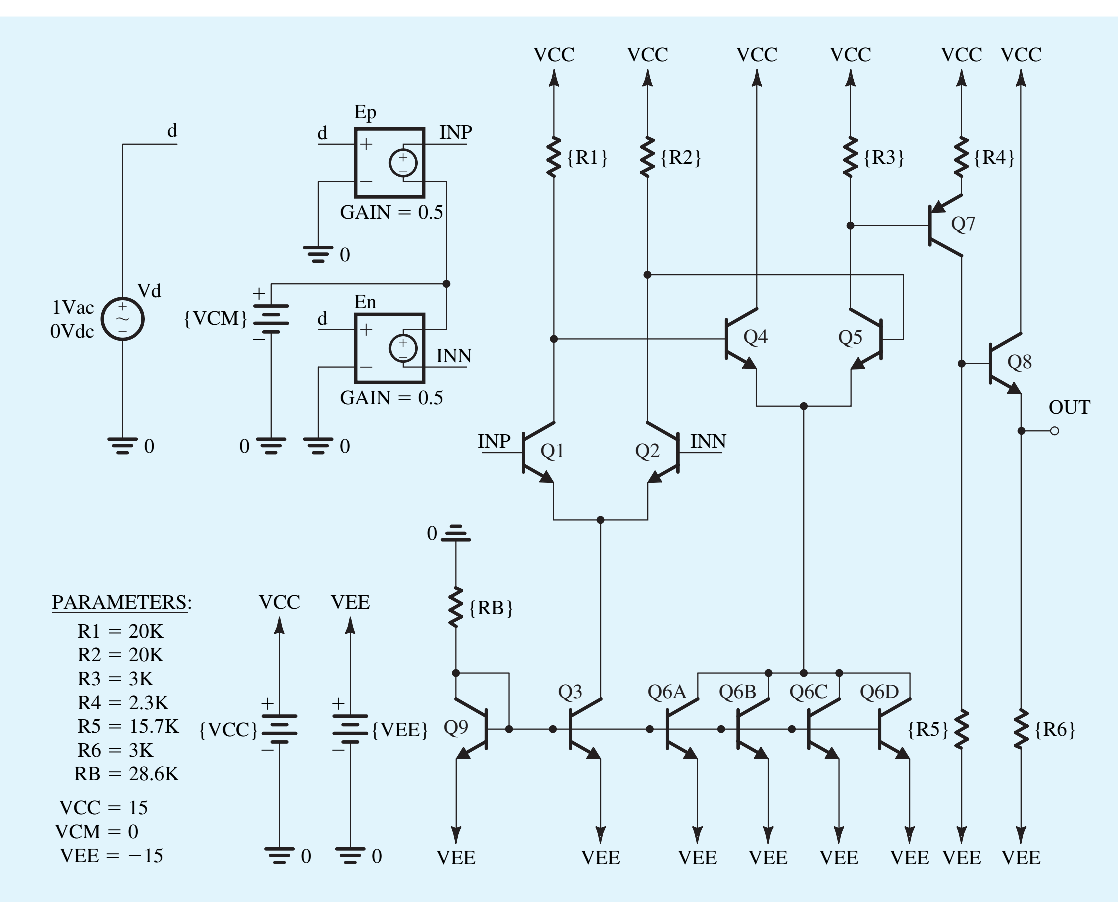


Figure B.26 Schematic capture of the op-amp circuit in Example 7.6.

To compute the **large-signal differential transfer characteristic** of the op-amp circuit, we perform a dc-analysis simulation in PSpice with the differential voltage input V_d swept over the range $-V_{EE}$ to $+V_{CC}$, and we plot the corresponding output voltage V_{OUT}. Figure B.27(a) shows the resulting dc transfer characteristic. The slope of this characteristic (i.e., dV_{OUT}/dV_d) corresponds to the differential gain of the amplifier. Note that, as expected, the high-gain region is in the vicinity of $V_d = 0\,\text{V}$. However, the resolution of the input-voltage axis is too coarse to yield much information about the details of the high-gain region. Therefore, to examine this region more closely, the dc analysis is repeated with V_d swept over the range $-5\,\text{mV}$ to $+5\,\text{mV}$ at increments of $10\,\mu\text{V}$. The resulting differential dc transfer characteristic is plotted in Fig. B.27(b). We observe that the linear region of the large-signal differential characteristic is bounded approximately by $V_d = -1.5\,\text{mV}$ and $V_d = +0.5\,\text{mV}$. Over this region, the output level changes from $V_{OUT} = -15\,\text{V}$ to about $V_{OUT} = +10\,\text{V}$ in a linear fashion. Thus, the output voltage swing for this amplifier is between $-15\,\text{V}$ and $+10\,\text{V}$, a rather asymmetrical range. A rough estimate for the differential gain of this amplifier can be obtained from the boundaries of the linear region as $A_d = [10-(-15)]\,\text{V}/[0.5-(-1.5)]\,\text{mV} = 12.5\times10^3\,\text{V/V}$. We also observe from Fig B.27(b) that $V_d \simeq -260\,\mu\text{V}$ when $V_{OUT} = 0$. Therefore, the amplifier has an input offset voltage V_{OS} of $+260\,\mu\text{V}$ (by convention, the negative value of the x-axis intercept of the large-signal differential transfer characteristic). This corresponds to an output offset voltage of $A_d V_{OS} \simeq \left(12.5\times10^3\right)(260\,\mu\text{V}) = 3.25\,\text{V}$, which is

Example PS.9.1 *continued*

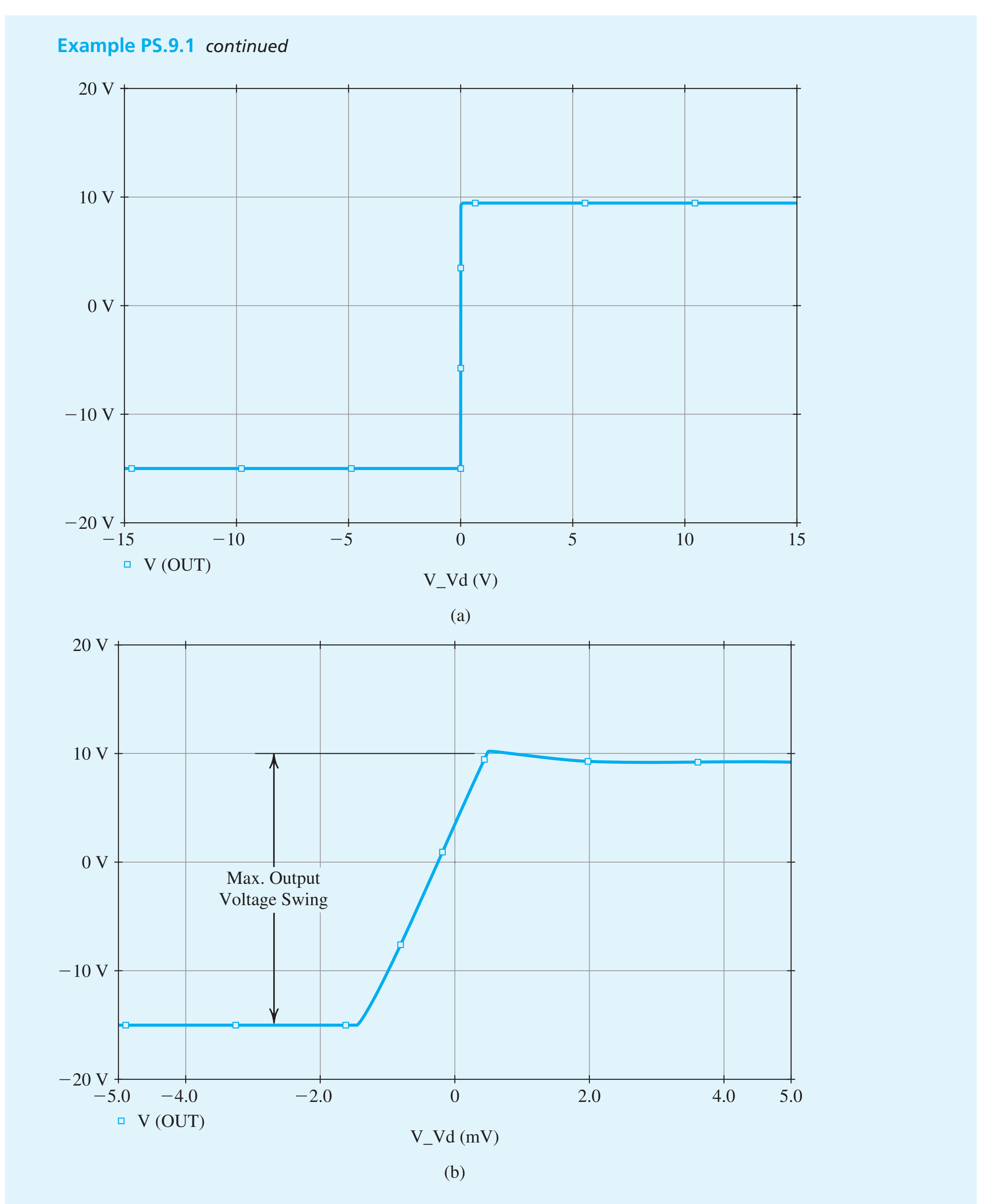


Figure B.27 **(a)** The large-signal differential transfer characteristic of the op-amp circuit in Fig. B.26. The common-mode input voltage V_{CM} is set to 0 V. **(b)** An expanded view of the transfer characteristic in the high-gain region.

close to the value found through the bias-point simulation. It should be emphasized that this offset voltage is inherent in the design and is not the result of component or device mismatches. Thus, it is usually referred to as a **systematic offset**.

Next, to compute the frequency response of the op-amp circuit[12] and to measure its differential gain A_d and its 3-dB frequency f_H in PSpice, we set the differential input voltage V_d to be a 1-V ac signal (with 0-V dc level), perform an ac-analysis simulation, and plot the output voltage magnitude $|V_{OUT}|$ versus frequency. Figure B.28(a) shows the resulting frequency response. Accordingly, $A_d = 13.96 \times 10^3$ V/V or 82.8 dB, and $f_H = 256.9$ kHz. Thus, this value of A_d is close to the value estimated using the large-signal differential transfer characteristic.

An approximate value of f_H can also be obtained using the expressions derived in Section 10.7. Specifically,

$$f_H \simeq \frac{1}{2\pi R_{eq} C_{eq}} \tag{B.26}$$

where

$$C_{eq} = C_{\mu2} + C_{\pi5} + C_{\mu5}\left[1 + g_{m5}\left(R_3 \| r_{o5} \| \left(r_{\pi7} + (\beta+1)R_4\right)\right)\right]$$

and

$$R_{eq} = R_2 | r_{o2} \| r_{\pi5}$$

The values of the small-signal parameters as computed by PSpice can be found in the output file of a bias-point (or an ac-analysis) simulation. Using these values results in $C_{eq} = 338$ pF, $R_{eq} = 2.91$ kΩ, and $f_H = 161.7$ kHz. However, this approximate value of f_H is much smaller than the value computed by PSpice. The reason for this disagreement is that the foregoing expression for f_H was derived (in Section 10.7) using the equivalent differential half-circuit concept. However, the concept is accurate only when it is applied to a symmetrical circuit. The op-amp circuit in Fig. B.26 is not symmetrical because the second gain stage formed by the differential pair Q_4–Q_5 has a load resistor R_3 in the collector of Q_5 only. To verify that the expression for f_H in Eq. (B.26) gives a close approximation for f_H in the case of a symmetric circuit, we insert a resistor R_3' (whose size is equal to R_3) in the collector of Q_4. Note that this will have only a minor effect on the dc operating point. The op-amp circuit with Q_4 having a collector resistor R_3' is then simulated in PSpice. Figure B.28(b) shows the resulting frequency response of this symmetric op amp, where $f_H = 155.7$ kHz. Accordingly, in the case of a perfectly symmetric op-amp circuit, the value of f_H in Eq. (B.26) closely approximates the value computed by PSpice. Comparing the frequency responses of the nonsymmetric (Fig. B.28a) and the symmetric (Fig. B.28b) op-amp circuits, we note that the 3-dB frequency of the op amp drops from 256.9 kHz to 155.7 kHz when resistor R_3' is inserted in the collector of Q_4 to make the op-amp circuit symmetrical. This is because, with a resistor R_3', the collector of Q_4 is no longer at signal ground and, hence, $C_{\mu4}$ experiences the Miller effect. Consequently, the high-frequency response of the op-amp circuit is degraded.

Observe that in the preceding ac-analysis simulation, owing to the systematic offset inherent in the design, the op-amp circuit is operating at an output dc voltage of 3.62 V. However, in an actual circuit implementation (with $V_{CM} = 0$), negative feedback is employed (see Chapters 2 and 11)

[12]This part of the example requires study of Section 10.7.

Example PS.9.1 *continued*

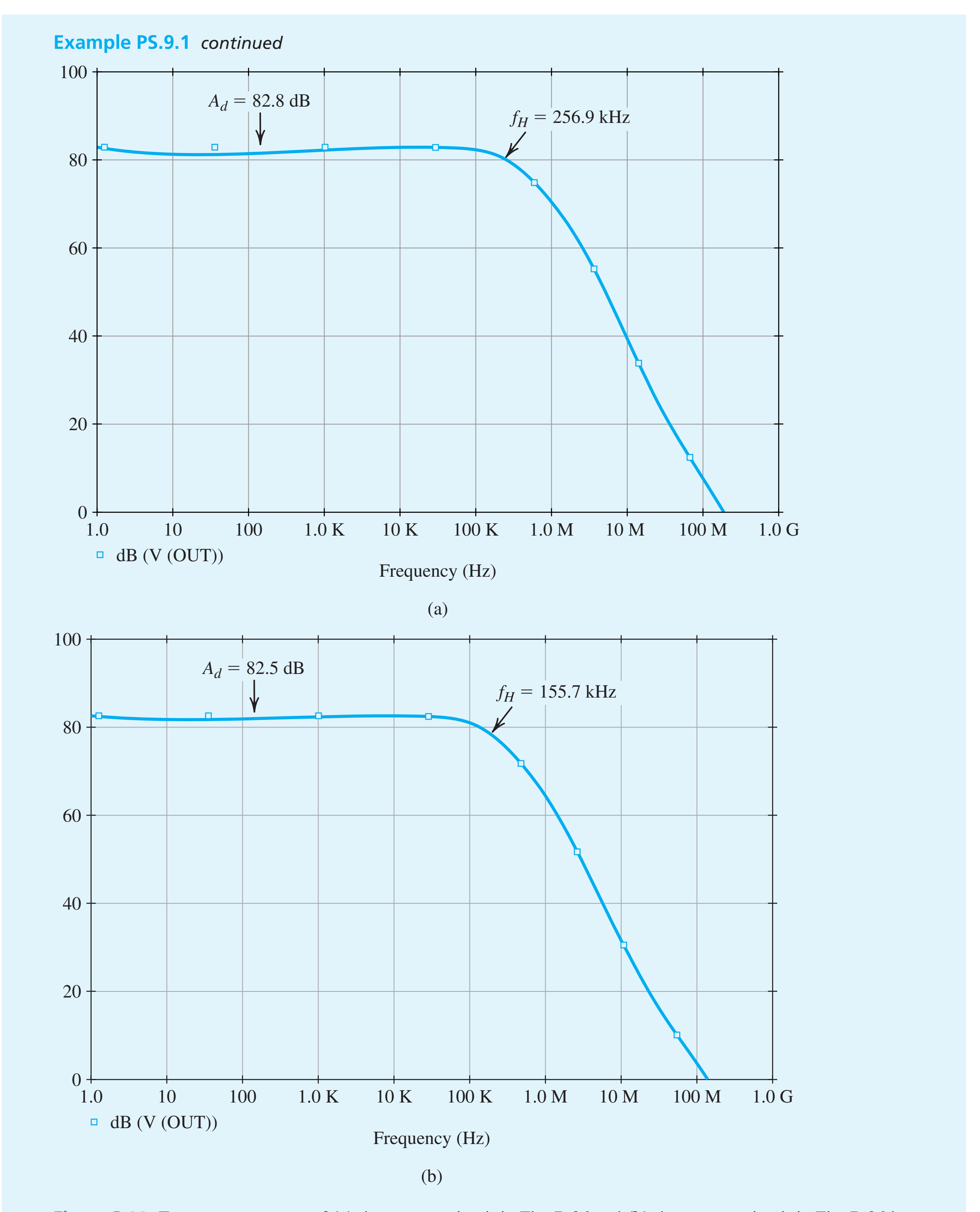
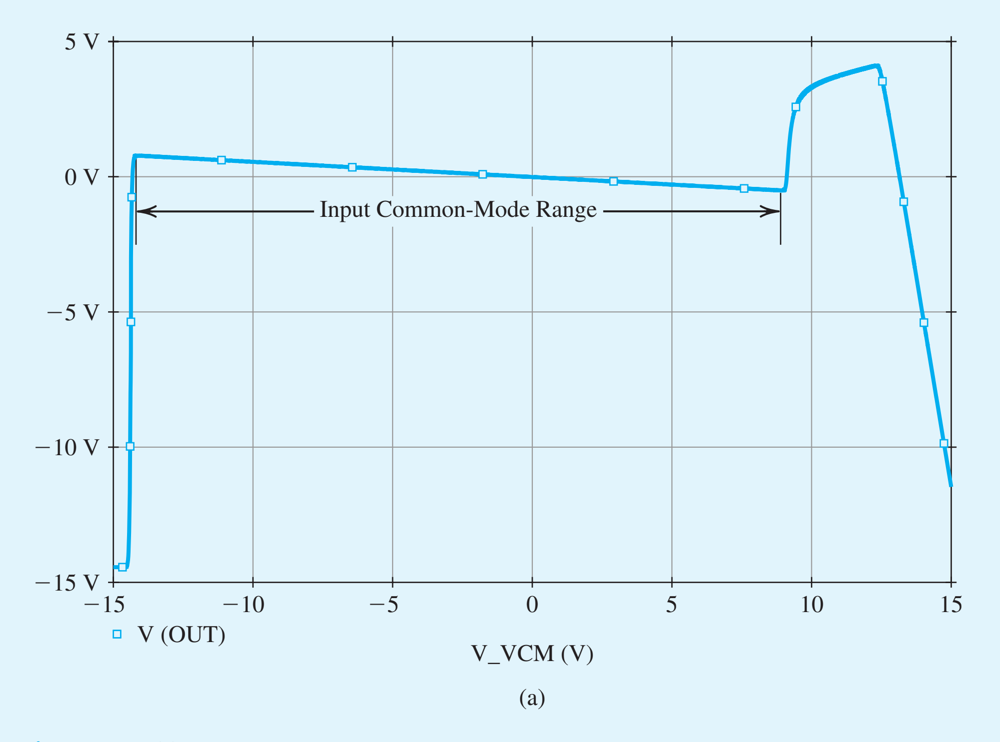


Figure B.28 Frequency response of **(a)** the op-amp circuit in Fig. B.26 and **(b)** the op-amp circuit in Fig. B.26 but with a resistor $R_3' = R_3$ inserted in the collector of Q_4 to make the op-amp circuit symmetrical.

and the output dc voltage is stabilized at zero. Thus, the small-signal performance of the op-amp circuit can be more accurately simulated by biasing the circuit so as to force operation at this level of output voltage. This can be easily done by applying a differential dc input of $-V_{OS}$. Superimposed on this dc input, we can apply an ac signal to perform an ac-analysis simulation for the purpose of, for example, computing the differential gain and the 3-dB frequency.

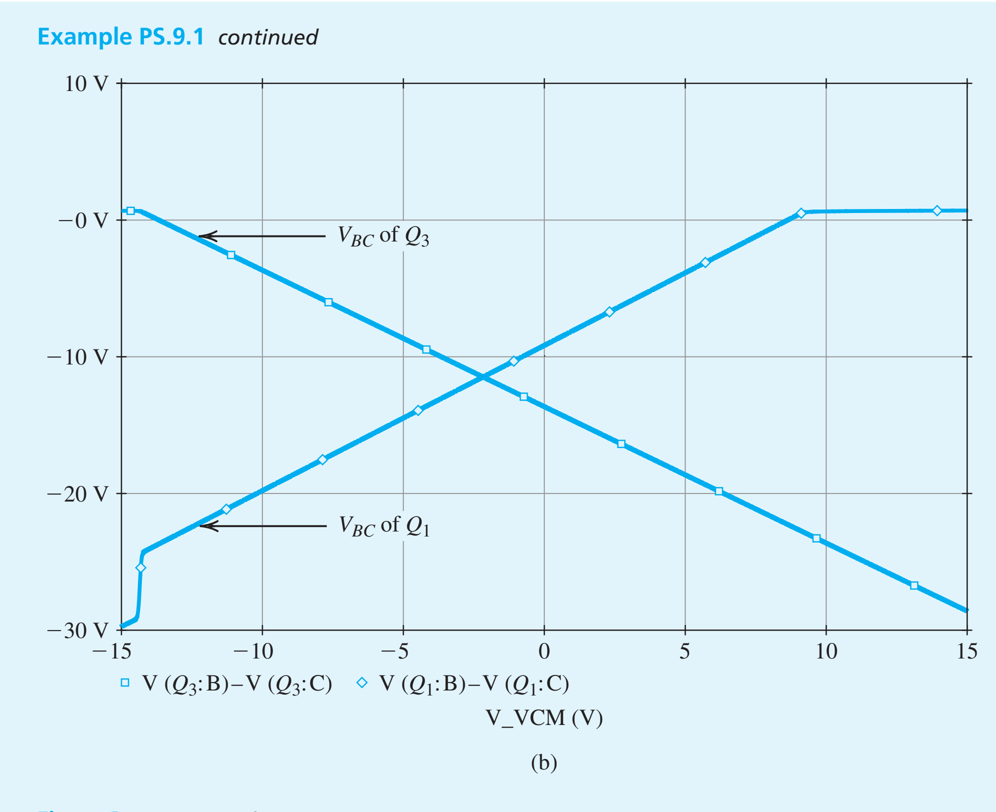
Finally, to compute the input common-mode range of the op-amp circuit in Fig. B.26, we perform a dc-analysis simulation in PSpice with the input common-mode voltage swept over the range $-V_{EE}$ to V_{CC}, while maintaining V_d constant at $-V_{OS}$ in order to cancel the output offset voltage (as discussed earlier) and, thus, prevent premature saturation of the BJTs. The corresponding output voltage V_{OUT} is plotted in Fig. B.29(a). From this common-mode dc transfer characteristic we find that the amplifier behaves linearly over the V_{CM} range –14.1 V to +8.9 V, which is therefore the **input common-mode range**. In Example 8.6, we noted that the upper limit of this range is determined by Q_1 and Q_2 saturating, whereas the lower limit is determined by Q_3 saturating. To verify this assertion, we requested PSpice to plot the values of the collector–base voltages of these BJTs versus the input common-mode voltage V_{CM}. The results are shown in Fig. B.29(b), from which we note that our assertion is indeed correct (recall that an *npn* BJT enters its saturation region when its base–collector junction becomes forward biased, i.e., $V_{BC} \geq 0$).



(a)

Figure B.29 **(a)** The large-signal common-mode transfer characteristic of the op-amp circuit in Fig. B.26. The differential input voltage V_d is set to $-V_{OS} = -260\,\mu\text{V}$ to prevent premature saturation. **(b)** The effect of the common-mode input voltage VCM o n the linearity of the input stage of the op-amp circuit in Fig. B.26. The base–collector voltage of Q_1 and Q_3 is shown as a function of VCM. The input stage of the op-amp circuit leaves the active region when the base–collector junction of either Q_1 or Q_3 becomes forward biased (i.e., when VBC ≥ 0).

Example PS.9.1 *continued*



(b)

Figure B.29 *continued*

Example PS.10.1

Frequency Response of the CMOS CS and the Folded-Cascode Amplifiers

In this example, we will use PSpice to compute the frequency response of both the CS and the folded-cascode amplifiers whose schematic capture diagrams are shown shortly in Figs. B.30 and B.32, respectively. We will assume that the dc bias levels at the output of the amplifiers are stabilized using negative feedback. However, before performing a small-signal analysis (an ac-analysis simulation) in SPICE to measure the frequency response, we will perform a dc analysis (a bias-point simulation) to verify that all MOSFETs are operating in the saturation region and, hence, ensure that the amplifier is operating in its linear region.

In the following, we will assume a 0.5-μm CMOS technology for the MOSFETs and use parts NMOSOP5 and PMOSOP5 whose SPICE level-1 model parameters are listed in Table B.3. To specify the dimensions of the MOSFETs in PSpice, we will use the multiplicative factor m, together with the channel length L and channel width W (as we did in Example PS.7.1).

The CMOS CS Amplifier

The CS amplifier circuit in Fig. B.30 is identical to the one shown in Fig. 7.4, except that a current source is connected to the source of the input transistor M_1 to set its drain current I_{D1} independently of its drain voltage V_{D1}. Furthermore, in our PSpice simulations, we used an impractically large bypass capacitor C_S of 1 F.

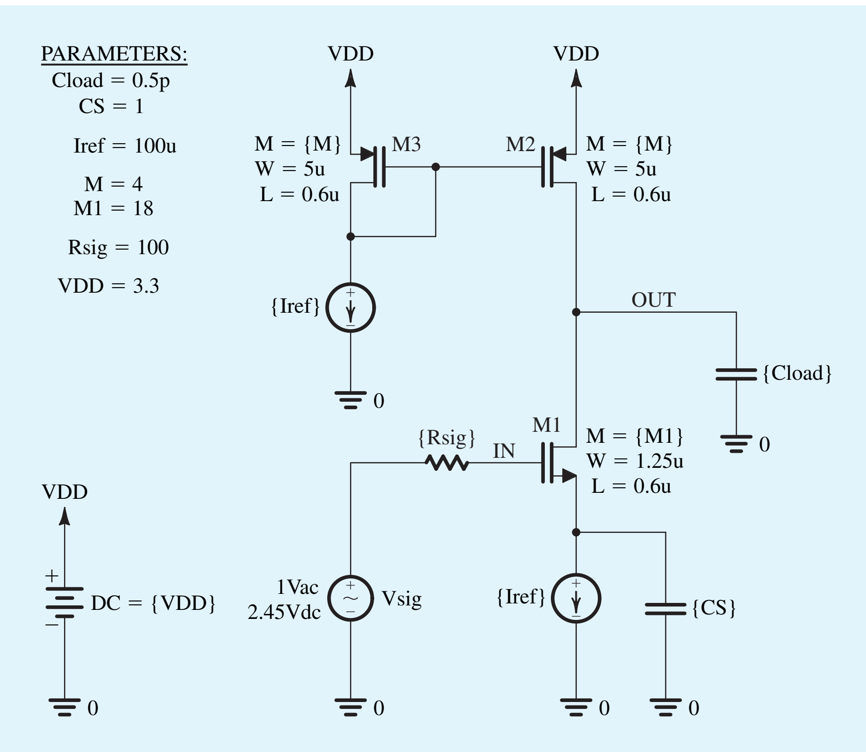


Figure B.30 Schematic capture of the CS amplifier in Example PS.10.1

This sets the source of M_1 at approximately signal ground during the ac-analysis simulation. Accordingly, the CS amplifier circuits in Figs. 7.4 and B.30 are equivalent for the purpose of frequency-response analysis. In Chapter 9, we found out, in the context of studying the differential pair, how the goals of this biasing approach for the CS amplifier are realized in practical IC implementations.

The CS amplifier in Fig. B.30 is designed assuming a reference current $I_{\text{ref}} = 100\,\mu\text{A}$ and $V_{DD} = 3.3$ V. The current-mirror transistors, M_2 and M_3, are sized for $V_{OV2} = V_{OV3} = 0.3$ V, while the input transistor M_1 is sized for $V_{OV1} = 0.15$ V. Unit-size transistors are used with $W/L = 1.25\ \mu\text{m}/0.6$ 0.6 μm for the NMOS devices and $W/L = 5\ \mu\text{m}/0.6\ \mu\text{m}$ for the PMOS devices. Thus, using the square law $I_D - V_{OV}$ of the MOSFET together with the 0.5-μm CMOS process parameters in Table B.4, we find $m_1 = 18$ and $m_2 = m_3 = 4$. Furthermore, Eq. (B.25) gives $G_v = -44.4$ V/V for the CS amplifier.

In the PSpice simulations of the CS amplifier in Fig. B.30, the dc bias voltage of the signal source is set such that the voltage at the source terminal of M_1 is $V_{S1} = 1.3$ V. This requires the dc level of V_{sig} to be $V_{OV1} + V_{tn1} + V_{S1} = 2.45$ V because $V_{tn1} \simeq 1$ V as a result of the body effect on M_1. The reasoning behind this choice of V_{S1} is that, in a practical circuit implementation, the current source that feeds the source of M_1 is realized using a cascode current mirror such as the one in Fig. 8.32. In this case, the minimum voltage required across the current source (i.e., the minimum V_{S1}) is $V_t + 2V_{OV} = 1.3$ V, assuming $V_{OV} = 0.3$ V for the current-mirror transistors.

A bias-point simulation is performed in PSpice to verify that all MOSFETs are biased in the saturation region. Next, to compute the frequency response of the amplifier, we set the ac voltage of the signal source to 1 V, perform an ac-analysis simulation, and plot the output voltage magnitude versus frequency. Figure B.31(a) shows the resulting frequency response for $R_{\text{sig}} = 100\,\Omega$ and $R_{\text{sig}} = 1$ MΩ. In both cases, a load capacitance of $C_{\text{load}} = 0.5$ pF is used. The corresponding values of the 3-dB frequency f_H of the amplifier are given in Table B.10.

Example PS.10.1 *continued*

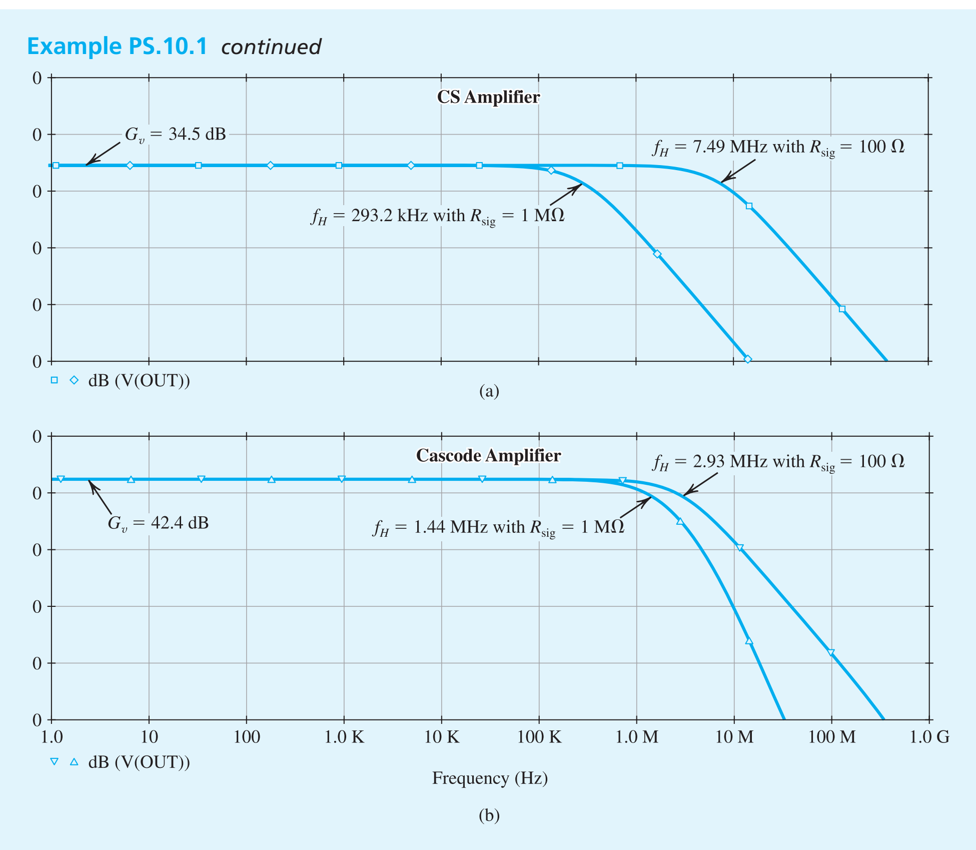


Figure B.31 Frequency response of **(a)** the CS amplifier and **(b)** the folded-cascode amplifier in Example PS.10.1, with $R_{sig} = 100\ \Omega$ and $R_{sig} = 1\ \text{M}\Omega$.

Observe that f_H drops when R_{sig} is increased. This is anticipated from our study of the high-frequency response of the CS amplifier in Section 10.3. Specifically, as R_{sig} increases, the pole

$$f_{p,\text{in}} = \frac{1}{2\pi} \frac{1}{R_{sig} C_{in}} \tag{B.27}$$

formed at the amplifier input will have an increasingly significant effect on the overall frequency response of the amplifier. As a result, the effective time constant τ_H in Eq. (10.88) increases and f_H decreases. When R_{sig} becomes very large, as it is when $R_{sig} = 1\ \text{M}\Omega$, a dominant pole is formed by R_{sig} and C_{in}. This results in

$$f_H \simeq f_{p,\text{in}} \tag{B.28}$$

To estimate $f_{p,\text{in}}$, we need to calculate the input capacitance C_{in} of the amplifier. Using Miller's theorem, we have

$$\begin{aligned} C_{in} &= C_{gs1} + C_{gd1}(1 + g_{m1} R'_L) \\ &= \left(\frac{2}{3} m_1 W_1 L_1 C_{ox} + C_{gs,ov1}\right) + C_{gd,ov1}(1 + g_{m1} R'_L) \end{aligned} \tag{B.29}$$

Table B.10 Dependence of the 3-dB Bandwidth f_H on R_{sig} for the CS and the Folded-Cascode Amplifiers in Example PS.10.1

	f_H	
R_{sig}	CS Amplifier	Folded-Cascode Amplifier
100 Ω	7.49 MHz	2.93 MHz
1 MΩ	293.2 kHz	1.44 MHz

where

$$R'_L = r_{o1} \| r_{o2} \quad \text{(B.30)}$$

Thus, C_{in} can be calculated using the values of C_{gs1} and C_{gd1}, which are computed by PSpice and can be found in the output file of the bias-point simulation. Alternatively, C_{in} can be found using Eq. (B.29) with the values of the overlap capacitances $C_{gs,ov1}$ and $C_{gsd,ov1}$ calculated using the process parameters in Table B.4 (as described in Eqs. B.9 and B.10); that is:

$$C_{gs,ov1} = m_1 W_1 CGSO \quad \text{(B.31)}$$

$$C_{gd,ov1} = m_1 W_1 CGDO \quad \text{(B.32)}$$

This results in $C_{in} = 0.53$ pF when $|G_v| = g_{m1}R'_L = 53.2$ V/V. Accordingly, using Eqs. (B.27) and (B.28), $f_H = 300.3$ kHz when $R_{sig} = 1$ MΩ, which is close to the value computed by PSpice.

The Folded-Cascode Amplifier

The folded-cascode amplifier circuit in Fig. B.32 is equivalent to the one in Fig. 8.16, except that a current source is placed in the source of the input transistor M_1 (for the same dc-biasing purpose as in the case of the CS amplifier). Note that, in Fig. B.32, the PMOS current mirror M_3–M_4 and the NMOS current mirror M_5–M_6 are used to realize, respectively, current sources I_1 and I_2 in the circuit of Fig. 8.16. Furthermore, the current transfer ratio of mirror M_3–M_4 is set to 2 (i.e., $m_3/m_4 = 2$). This results in $I_{D3} \simeq 2I_{ref}$. Hence, transistor M_2 is biased at $I_{D2} = I_{D3} - I_{D1} = I_{ref}$. The gate bias voltage of transistor M_2 is generated using the diode-connected transistors M_7 and M_8. The size and drain current of these transistors are set equal to those of transistor M_2. Therefore, ignoring the body effect,

$$V_{G2} = V_{DD} - V_{SG7} - V_{SG8} \simeq V_{DD} - 2(|V_{tp}| + |V_{OVp}|)$$

where V_{OVp} is the overdrive voltage of the PMOS transistors in the amplifier circuit. These transistors have the same overdrive voltage because their I_D/m is the same. Thus, such a biasing configuration results in $V_{SG2} = |V_{tp}| + |V_{OVp}|$ as desired, while setting $V_{SD3} = |V_{tp}| + |V_{OVp}|$ to improve the bias matching between M_3 and M_4.

The folded-cascode amplifier in Fig. B.32 is designed assuming a reference current $I_{ref} = 100$ µA and $V_{DD} = 3.3$ V (similar to the case of the CS amplifier). All transistors are sized for an overdrive voltage of 0.3 V, except for the input transistor M_1, which is sized for $V_{OV1} = 0.15$ V. Thus, since $I_D = \frac{1}{2}\mu m C_{ox} m (W/L_{eff})V_{OV}^2$, all the MOSFETs in the amplifier circuit are designed using $m = 4$, except for $m_1 = 18$.

The midband voltage gain of the folded-cascode amplifier in Fig. B.32 can be expressed as

$$G_v = -g_{m1}R_{out} \quad \text{(B.33)}$$

where

$$R_{out} = R_{out2} \| R_{out5} \quad \text{(B.34)}$$

Example PS.10.1 *continued*

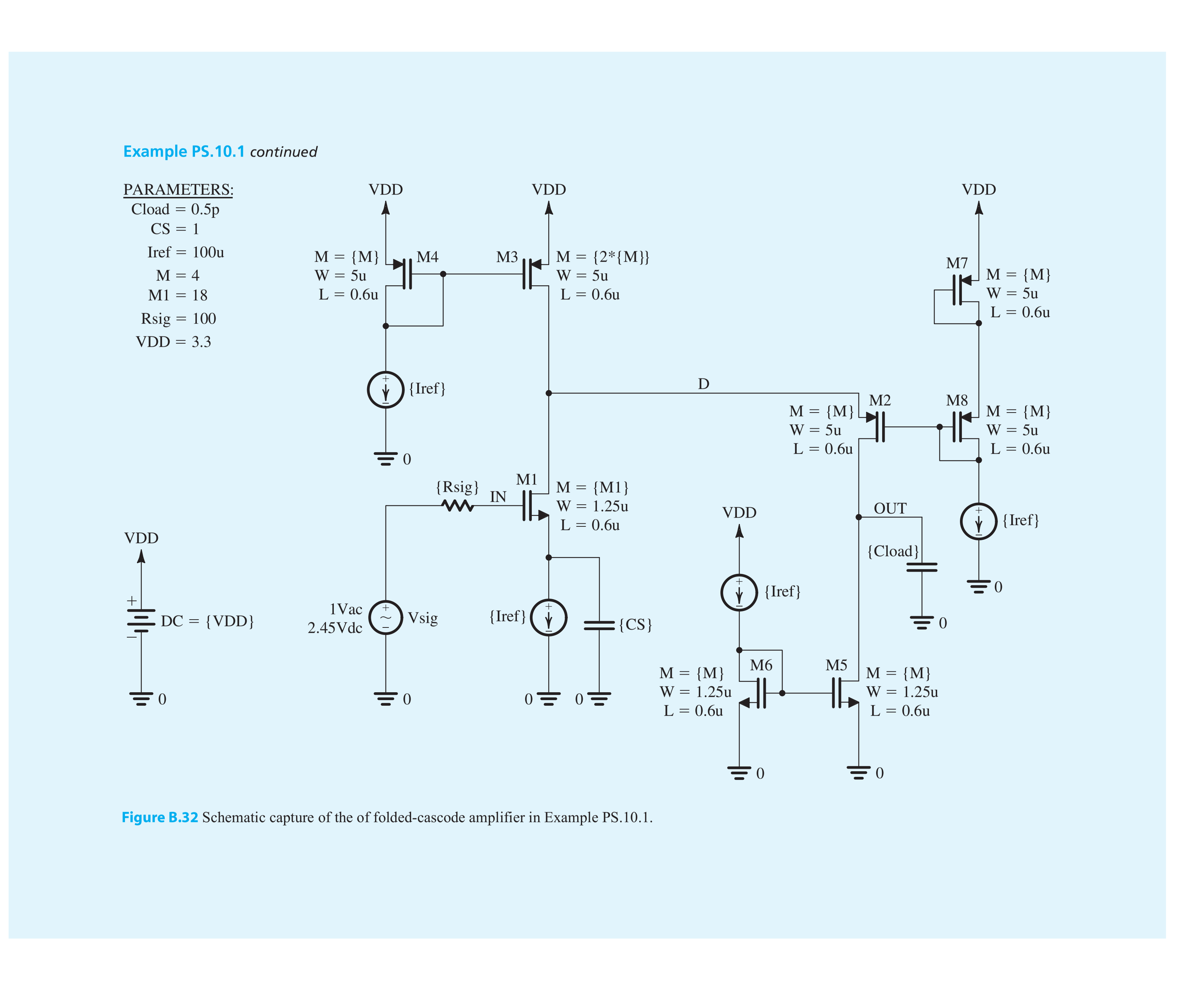


Figure B.32 Schematic capture of the of folded-cascode amplifier in Example PS.10.1.

is the output resistance of the amplifier. Here, R_{out2} is the resistance seen looking into the drain of the cascode transistor M_2, while R_{out5} is the resistance seen looking into the drain of the current-mirror transistor M_5. Using Eq. (8.72), we have

$$R_{out2} \simeq (g_{m2}r_{o2})R_{s2} \tag{B.35}$$

where

$$R_{s2} = r_{o1} \parallel r_{o3} \tag{B.36}$$

is the effective resistance at the source of M_2. Furthermore,

$$R_{out5} = r_{o5} \tag{B.37}$$

Thus, for the folded-cascoded amplifier in Fig. B.32,

$$R_{out} \simeq r_{o5} \tag{B.38}$$

and

$$G_v \simeq -g_{m1}r_{o5} = -2\frac{V_{An}}{V_{OV1}} \tag{B.39}$$

Using the 0.5-μm CMOS parameters, this gives $R_{out} = 100$ kΩ and $G_v = -133$ V/V. Therefore, R_{out} and hence $|G_v|$ of the folded-cascode amplifier in Fig. B.32 are larger than those of the CS amplifier in Fig. B.30 by a factor of 3.

Figure B.31(b) shows the frequency response of the folded-cascode amplifier as computed by PSpice for the cases of $R_{sig} = 100\ \Omega$ and $R_{sig} = 1\ \text{M}\Omega$. The corresponding values of the 3-dB frequency f_H of the amplifier are given in Table B.10. Observe that when R_{sig} is small, f_H of the folded-cascode amplifier is lower than that of the CS amplifier by a factor of approximately 2.6, approximately equal to the factor by which the gain is increased. This is because when R_{sig} is small, the frequency response of both amplifiers is dominated by the pole formed at the output node, that is,

$$f_H \simeq f_{p,out} = \frac{1}{2\pi}\frac{1}{R_{out}C_{out}} \tag{B.40}$$

Since the output resistance of the folded-cascode amplifier is larger than that of the CS amplifier (by a factor of approximately 3, as found through the hand analysis above) while their output capacitances are approximately equal, the folded-cascode amplifier has a lower f_H in this case.

On the other hand, when R_{sig} is large, f_H of the folded-cascode amplifier is much higher than that of the CS amplifier. This is because, in this case, the effect of the pole at $f_{p,in}$ on the overall frequency response of the amplifier becomes significant. Since, due to the Miller effect, C_{in} of the CS amplifier is much larger than that of the folded-cascode amplifier, its f_H is much lower in this case. To confirm this point, observe that C_{in} of the folded-cascode amplifier can be estimated by replacing R_L' in Eq. (B.29) with the total resistance R_{d1} between the drain of M_1 and ground. Here,

$$R_{d1} = r_{o1} \parallel r_{o3} \parallel R_{in2} \tag{B.41}$$

where R_{in2} is the input resistance of the common-gate transistor M_2 and can be obtained using an approximation of the relationship in Eq. (8.53) as

$$R_{in2} \simeq \frac{r_{o2} + r_{o5}}{g_{m2}r_{o2}} \tag{B.42}$$

Thus,

$$R_{d1} \simeq r_{o1} \parallel r_{o3} \parallel \left(\frac{r_{o2} + r_{o5}}{g_{m2}r_{o2}} \simeq \frac{2}{g_{m2}}\right) \tag{B.43}$$

Example PS.10.1 *continued*

Therefore, R_{d1} is much smaller than R_L' in Eq. (B.30). Hence, C_{in} of the folded-cascode amplifier in Fig. B.32 is indeed much smaller than that of the CS amplifier in Fig. B.30. This confirms that the folded-cascode amplifier is much less impacted by the Miller effect and, therefore, can achieve a much higher f_H when R_{sig} is large.

The midband gain of the folded-cascode amplifier can be significantly increased by replacing the current mirror M_5–M_6 with a current mirror having a larger output resistance, such as the cascode current mirror in Fig. 8.32 whose output resistance is approximately $g_m r_o^2$. In this case, however, R_{in2} and hence R_{d1} increase, causing an increased Miller effect and a corresponding reduction in f_H.

Finally, it is interesting to observe that the frequency response of the folded-cascode amplifier, shown in Fig. B.31(b), drops beyond f_H at approximately –20 dB/decade when $R_{sig} = 100\,\Omega$ and at approximately −40 dB/decade when $R_{sig} = 1\,\mathrm{M}\Omega$. This is because when R_{sig} is small, the frequency response is dominated by the pole at $f_{p,out}$. However, when R_{sig} is increased, $f_{p,in}$ is moved closer to $f_{p,out}$ and both poles contribute to the gain falloff.

Example PS.11.1

Determining the Loop Gain of a Feedback Amplifier

This example illustrates the use of SPICE to compute the loop gain $A\beta$. For this purpose, we shall use the shunt–series feedback amplifier shown in Fig. B.33a.

To compute the loop gain, we set the input signal V_s to zero, and we choose to break the feedback loop between the collector of Q_1 and the base of Q_2. However, in breaking the feedback loop, we must ensure that the following two conditions that existed prior to breaking the feedback loop do not change: (1) the dc bias situation and (2) the ac signal termination.

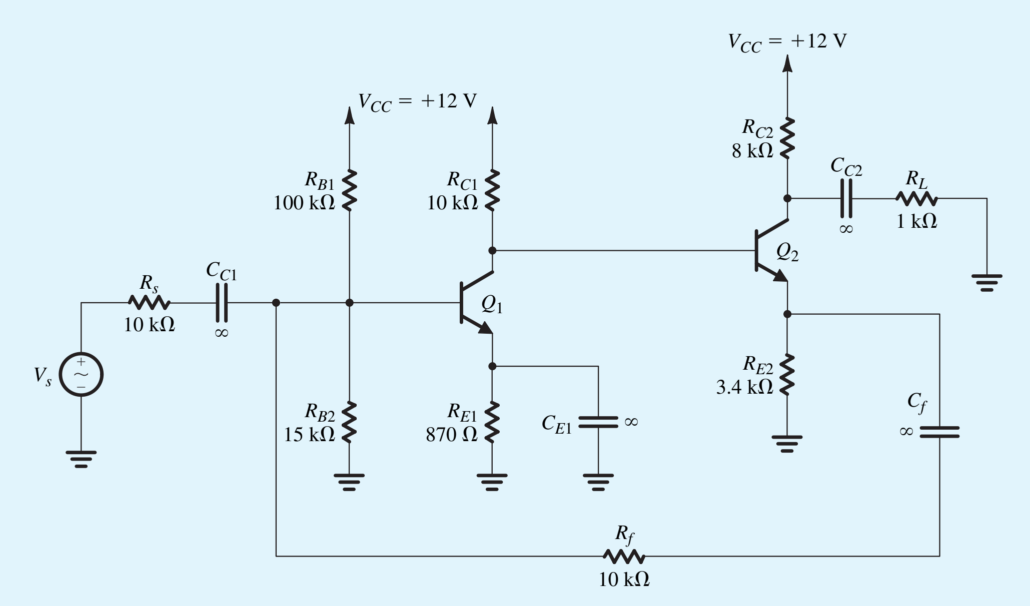


Figure B.33 Circuit of the shunt–series feedback amplifier in Example PS.11.1.

To break the feedback loop without disturbing the dc bias conditions of the circuit, we insert a large inductor L_{break}, as shown in Fig. B.34(a). Using a value of, say, $L_{\text{break}} = 1$ GH will ensure that the loop is opened for ac signals while keeping dc bias conditions unchanged.

To break the feedback loop without disturbing the signal termination conditions, we must load the loop output at the collector of Q_1 with a termination impedance Z_t whose value is equal to the impedance seen looking into the loop input at the base of Q_2. Furthermore, to avoid disturbing the dc bias conditions, Z_t must be connected to the collector of Q_1 via a large coupling capacitor. However, it is not always easy to determine the value of the termination impedance Z_t. So, we will describe two simulation methods to compute the loop gain without explicitly determining Z_T.

Method 1 *Using the open-circuit and short-circuit transfer functions*

As described in Section 11.1, the loop gain can be expressed as

$$A\beta = -1 \Big/ \left(\frac{1}{T_{oc}} + \frac{1}{T_{sc}} \right)$$

where T_{oc} is the open-circuit voltage transfer function and T_{sc} is the short-circuit voltage transfer function.

The circuit for determining T_{oc} is shown in Fig. B.34(b). Here, an ac test signal voltage V_t is applied to the loop input at the base of Q_2 via a large coupling capacitor (having a value of, say, 1 kF) to avoid disturbing the dc bias conditions. Then,

$$T_{oc} = \frac{V_{oc}}{V_t}$$

where V_{oc} is the ac open-circuit output voltage at the collector of Q_1.

In the circuit for determining T_{sc} (Fig. B.34b), an ac test signal current I_t is applied to the loop input at the base of Q_2. Note that a coupling capacitor is not needed in this case because the ac current source appears as an open circuit at dc, and, hence, does not disturb the dc bias conditions.

The loop output at the collector of Q_1 is ac short-circuited to ground via a large capacitor C_{to}. Then,

$$T_{sc} = \frac{I_{sc}}{I_t}$$

where I_{sc} is the ac short-circuit output current at the collector of Q_1.

Method 2 *Using a replica circuit*

As shown in Fig. B.35, a replica of the feedback amplifier circuit can be simply used as a termination impedance. Here, the feedback loops of both the amplifier circuit and the replica circuit are broken using a large inductor L_{break} to avoid disturbing the dc bias conditions. The loop output at the collector of Q_1 in the amplifier circuit is then connected to the loop input at the base of Q_2 in the replica circuit via a large coupling capacitor C_{to} (again, to avoid disturbing the dc bias conditions). Thus, for ac signals, the loop output at the collector of Q_1 in the amplifier circuit sees an impedance equal to that seen before the feedback loop is broken. Accordingly, we have ensured that the conditions that existed in the amplifier circuit prior to breaking the loop have not changed.

Next, to determine the loop gain $A\beta$, we apply an ac test signal voltage V_t via a large coupling capacitor C_{ti} to the loop input at the base of Q_2 in the amplifier circuit. Then, as described in Section 11.5,

$$A\beta = -\frac{V_r}{V_t}$$

where V_r is the ac returned signal at the loop output at the collector of Q_1 in the amplifier circuit.

Example PS.11.1 *continued*

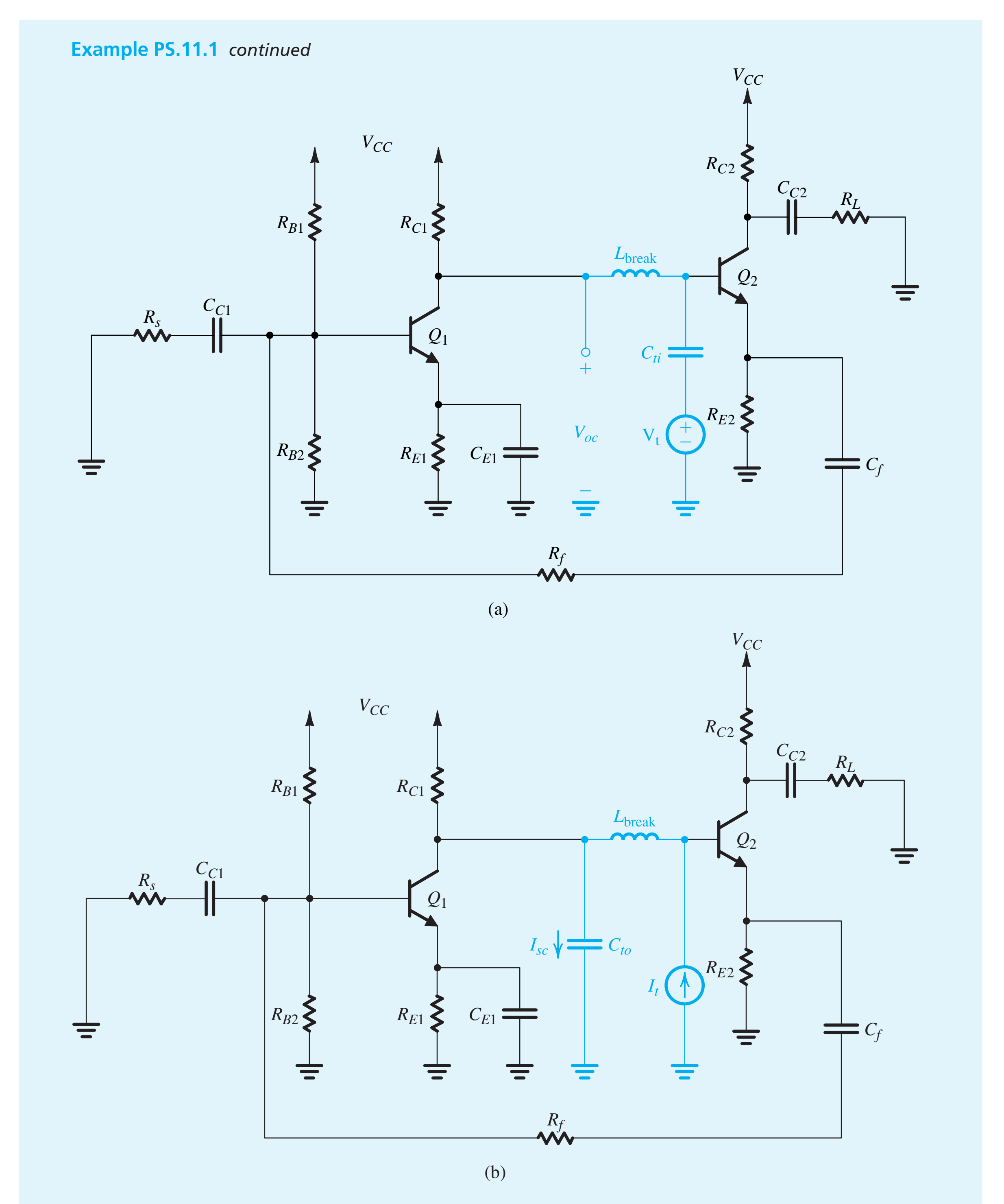


Figure B.34 Circuits for simulating **(a)** the open-circuit voltage transfer function T_{oc} and **(b)** the short-circuit current transfer function T_{sc} of the feedback amplifier in Fig. B.33 for the purpose of computing its loop gain.

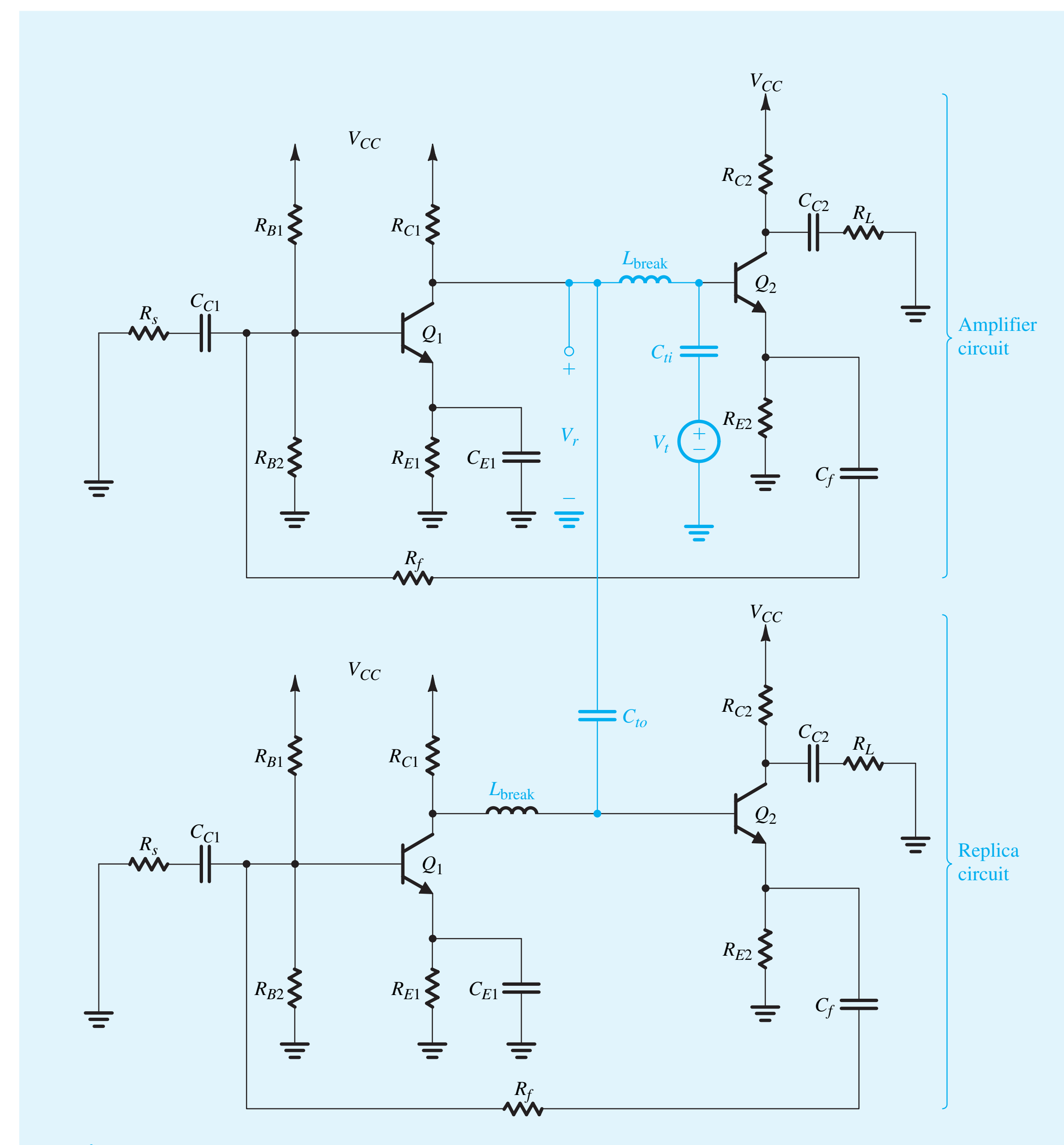


Figure B.35 Circuit for simulating the loop gain of the feedback amplifier circuit in Fig. B.33 using the replica-circuit method.

To compute the loop gain $A\beta$ of the feedback amplifier circuit in Fig. B.33 using PSpice, we choose to simulate the circuit in Fig. B.35. In the PSpice simulations, we used part Q2N3904 (whose SPICE model is given in Table B.6) for the BJTs, and we set L_{break} to be 1 GH and the coupling and bypass capacitors to be 1 kF. The magnitude and phase of $A\beta$ are plotted in Fig. B.36, from which we see that the feedback amplifier has a gain margin of 53.7 dB and a phase margin of 88.7°.

Example PS.11.1 *continued*

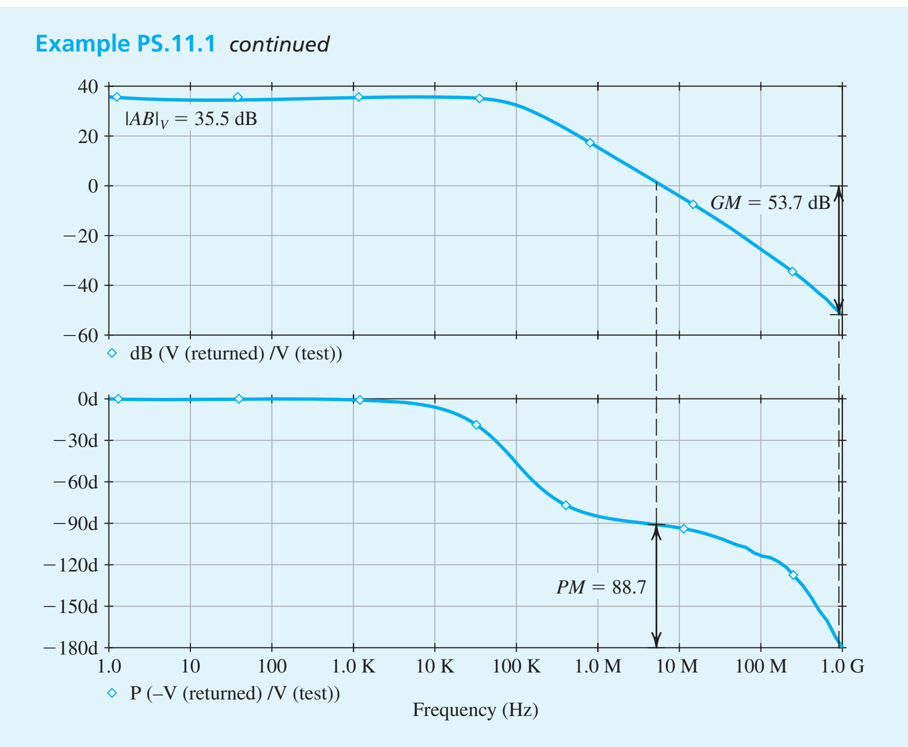


Figure B.36 **(a)** Magnitude and **(b)** phase of the loop gain $A\beta$ of the feedback amplifier circuit in Fig. B.33.

Example PS.12.1

Class B BJT Output Stage

We investigate the operation of the class B output stage whose schematic capture is shown in Fig. B.37. For the power transistors, we use the discrete BJTs MJE243 and MJE253 (from ON Semiconductor),[13] which are rated for a maximum continuous collector current $I_{C\max} = 4$ A and a maximum collector–emitter voltage of $V_{CE\max} = 100$ V. To permit comparison with the hand analysis performed in Example 11.1, in the simulation, we use component and voltage values identical (or close) to those of the circuit designed in Example 11.1. Specifically, we use a load resistance of 8 Ω, an input sine-wave signal of 17.9-V peak and 1-kHz frequency, and 23-V power supplies. In PSpice, a transient-analysis simulation is performed over the interval 0 ms to 3 ms, and the waveforms of various node voltages and branch currents are plotted. In this example, Probe (the graphical interface of PSpice) is utilized to compute various power-dissipation values. Some of the resulting waveforms are displayed in Fig. B.38. The upper and middle graphs show the load voltage and current, respectively. The peak voltage amplitude is 16.9 V, and the peak current amplitude is 2.1 A. If one looks carefully, one can observe that both exhibit crossover distortion. The bottom

[13] In PSpice, we have created BJT parts for these power transistors based on the values of the SPICE model parameters available on the data sheets available from ON Semiconductor. Readers can find these parts (labeled QMJE243 and QMJE253) in the SEDRA.olb library, which is available at www.oup.com/us/sedrasmith.

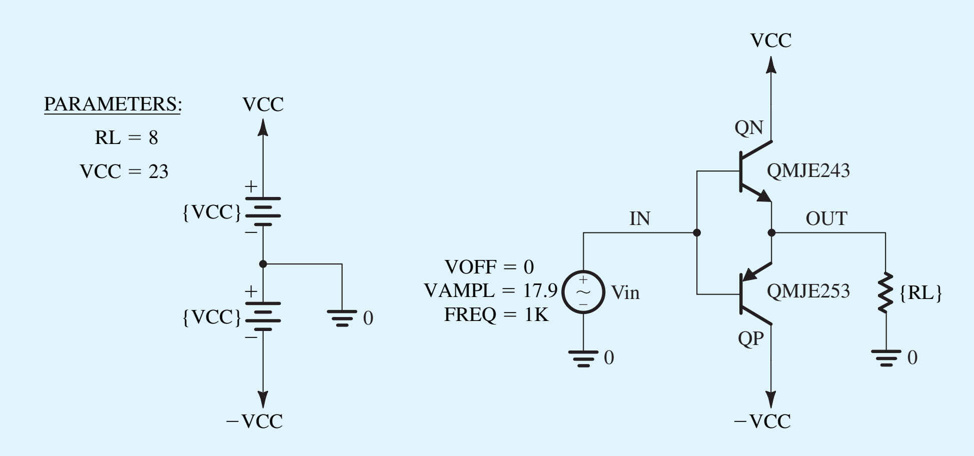


Figure B.37 Capture schematic of the class B output stage in Example PS.12.1.

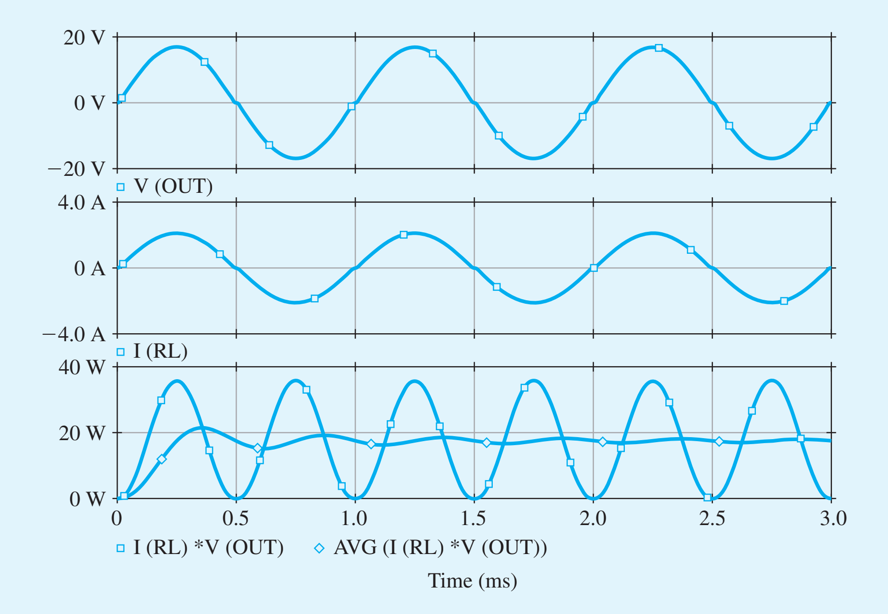


Figure B.38 Several waveforms associated with the class B output stage (shown in Fig. B.37) when excited by a 17.9-V, 1-kHz sinusoidal signal. The upper graph displays the voltage across the load resistance, the middle graph displays the load current, and the lower graph displays the instantaneous and average power dissipated by the load.

graph displays the instantaneous and the average power dissipated in the load resistance as computed using Probe by multiplying the voltage and current values to obtain the instantaneous power, and taking a running average for the average load power P_L. The transient behavior of the average load power, which eventually settles into a quasi-constant steady state of about 17.6 W, is an artifact of the PSpice algorithm used to compute the running average of a waveform.

The upper two graphs of Fig. B.39 show the voltage and current waveforms, respectively, of the positive supply, $+V_{CC}$. The bottom graph shows the instantaneous and average power supplied by $+V_{CC}$.

Example PS.12.1 *continued*

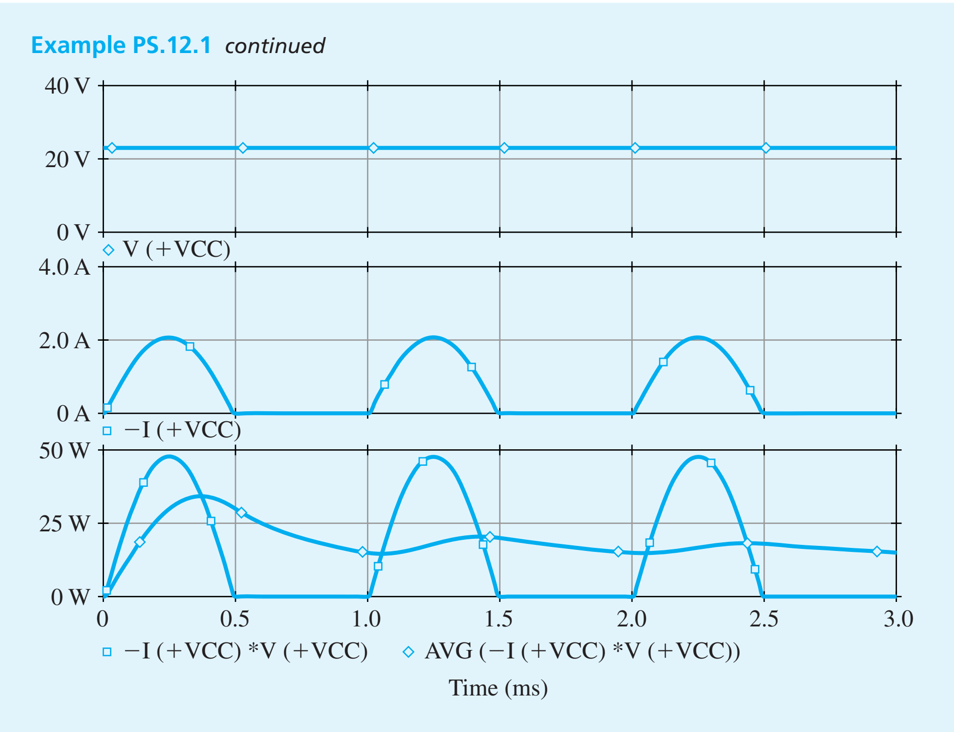


Figure B.39 The voltage (upper graph), current (middle graph), and instantaneous and average power (bottom graph) supplied by the positive voltage supply ($+V_{CC}$) in the circuit of Fig. B.37.

Similar waveforms can be plotted for the negative supply, $-V_{CC}$. The average power provided by each supply is found to be about 15 W, for a total supply power P_S of 30 W. Thus, the power-conversion efficiency can be computed to be

$$\eta = P_L/P_S = \frac{17.6}{30} \times 100\% = 58.6\%$$

Figure B.40 shows plots of the voltage, current, and power waveforms associated with transistor Q_P. Similar waveforms can be obtained for Q_N. As expected, the voltage waveform is a sinusoid, and the current waveform consists of half-sinusoids. The waveform of the instantaneous power, however, is rather unusual. It indicates the presence of some distortion as a result of driving the transistors rather hard. This can be verified by reducing the amplitude of the input signal. Specifically, when the amplitude is reduced to about 17 V, the “dip” in the power waveform vanishes. The average power dissipated in each of Q_N and Q_P can be computed by Probe and are found to be approximately 6 W.

Table B.11 provides a comparison of the results found from the PSpice simulation and the corresponding values obtained using hand analysis in Example 11.1. Observe that the two sets of results are quite close.

To investigate the crossover distortion further, we present in Fig. B.41 a plot of the voltage transfer characteristic (VTC) of the class B output stage. This plot is obtained through a dc-analysis simulation with v_{IN} swept over the range –10 V to +10 V in 1.0-mV increments. Using Probe, we determine that the slope of the VTC is nearly unity and that the dead band extends from –0.60 V to +0.58 V. The effect of the crossover distortion can be quantified by performing a Fourier analysis on the output voltage waveform in PSpice. This analysis decomposes the waveform generated through a transient analysis into its

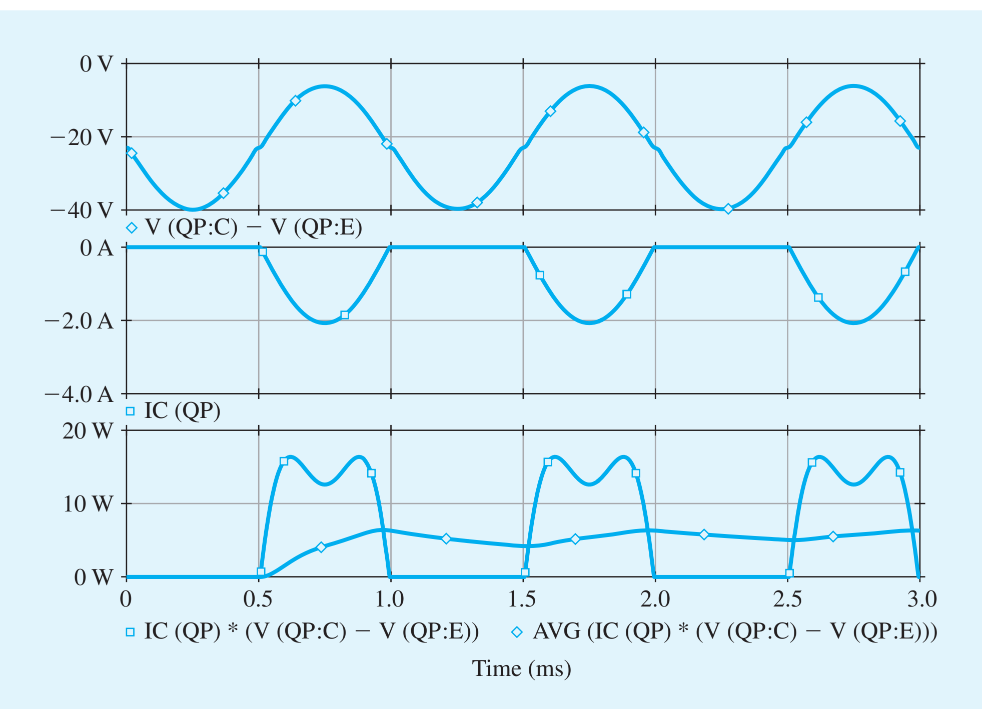


Figure B.40 Waveforms of the voltage across, the current through, and the power dissipated in the *pnp* transistor Q_P of the output stage shown in Fig. B.37.

Table B.11 Various Power Terms Associated with the Class B Output Stage Shown in Fig. B.37 as Computed by Hand and by PSpice Analysis

Power/Efficiency	Equation	Hand Analysis (Example PS.12.1)	PSpice	Error %[1]
P_S	$\frac{2}{\pi}\frac{\hat{V}_o}{R_L}V_{CC}$	31.2 W	30.0 W	4
P_D	$\frac{2}{\pi}\frac{\hat{V}_o}{R_L}V_{CC} - \frac{1}{2}\frac{\hat{V}_o^2}{R_L}$	13.0 W	12.4 W	4.6
P_L	$\frac{1}{2}\frac{\hat{V}_o^2}{R_L}$	18.2 W	17.6 W	3.3
η	$\frac{P_L}{P_S} \times 100\%$	58.3%	58.6%	–0.5

[1]Relative percentage error between the values predicted by hand and by PSpice.

Fourier-series components. Further, PSpice computes the total harmonic distortion (THD) of the output waveform. The results obtained from the simulation output file are shown on the next page.

These Fourier components are used to plot the line spectrum shown in Fig. B.42. We note that the output waveform is rather rich in odd harmonics and that the resulting THD is rather high (2.14%).

Example PS.12.1 *continued*

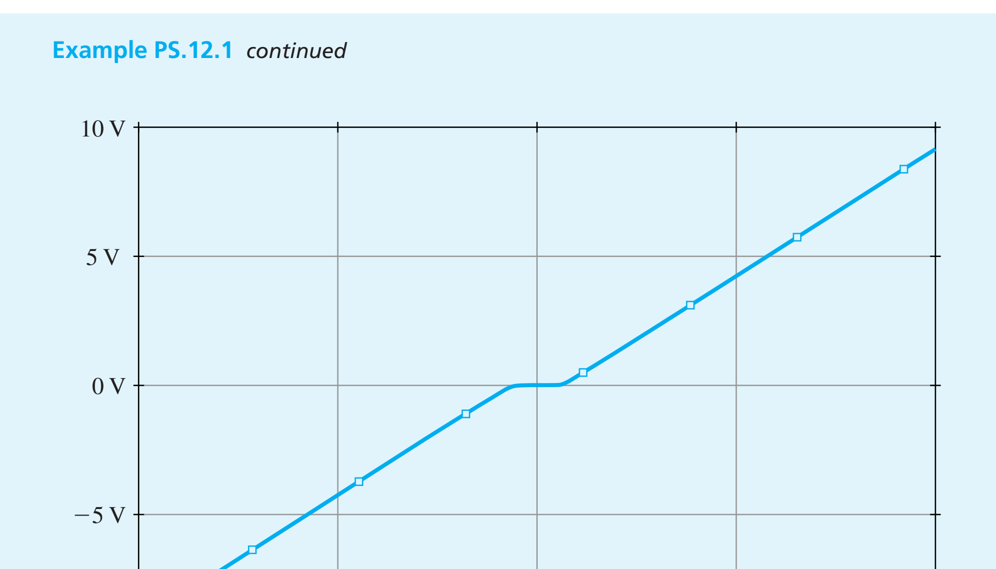


Figure B.41 Transfer characteristic of the class B output stage of Fig. B.37.

```
FOURIER COMPONENTS OF TRANSIENT RESPONSE V(OUT)

DC COMPONENT = -1.525229E-02

HARMONIC   FREQUENCY   FOURIER     NORMALIZED   PHASE       NORMALIZED
   NO        (HZ)      COMPONENT   COMPONENT    (DEG)       PHASE (DEG)

    1      1.000E+03   1.674E+01   1.000E+00   -2.292E-03   0.000E+00
    2      2.000E+03   9.088E-03   5.428E-04    9.044E+01   9.044E+01
    3      3.000E+03   2.747E-01   1.641E-02   -1.799E+02  -1.799E+02
    4      4.000E+03   4.074E-03   2.433E-04    9.035E+01   9.036E+01
    5      5.000E+03   1.739E-01   1.039E-02   -1.799E+02  -1.799E+02
    6      6.000E+03   5.833E-04   3.484E-05    9.159E+01   9.161E+01
    7      7.000E+03   1.195E-01   7.140E-03   -1.800E+02  -1.799E+02
    8      8.000E+03   5.750E-04   3.435E-05    9.128E+01   9.129E+01
    9      9.000E+03   9.090E-02   5.429E-03   -1.800E+02  -1.799E+02
   10      1.000E+04   3.243E-04   1.937E-05    9.120E+01   9.122E+01

TOTAL HARMONIC DISTORTION = 2.140017E+00 PERCENT
```

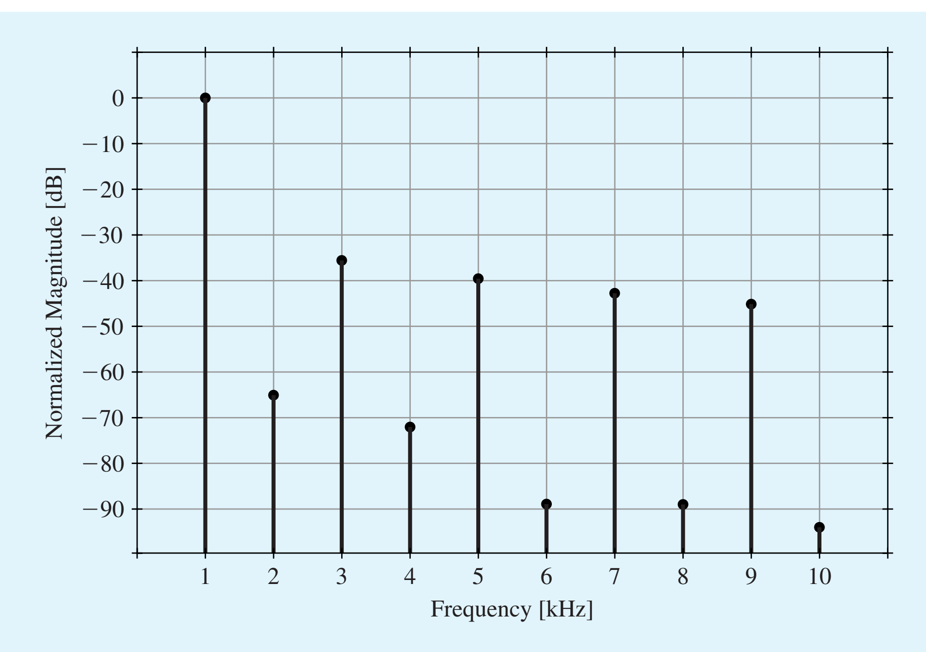


Figure B.42 Fourier-series components of the output waveform of the class B output stage in Fig. B.37.

Example PS.13.1

Frequency Compensation of the Two-Stage CMOS Op Amp

In this example, we will use PSpice to aid in designing the frequency compensation of the two-stage CMOS circuit whose capture schematic is shown in Fig. B.43. PSpice will then be employed to determine the frequency response and the slew rate of the op amp. We will assume a 0.5-μm *n*-well CMOS technology for the MOSFETs and will use the SPICE level-1 model parameters listed in Table B.4. Observe that to eliminate the body effect and improve the matching between M_1 and M_2, the source terminals of the input PMOS transistors M_1 and M_2 are connected to their *n* well.

The op-amp circuit in Fig. B.43 is designed using a reference current $I_{REF} = 90\ \mu A$, a supply voltage $V_{DD} = 3.3$ V, and a load capacitor $C_L = 1$ pF. Unit-size transistors with $W/L = 1.25\ \mu m/0.6\ \mu m$ are used for both the NMOS and PMOS devices. The transistors are sized for an overdrive voltage $V_{OV} = 0.3$V. The corresponding multiplicative factors are given in Fig. B.43.

In PSpice, the common-mode input voltage V_{CM} of the op-amp circuit is set to $V_{DD}/2 = 1.65$ V. A bias-point simulation is performed to determine the dc operating point. Using the values found in the simulation output file for the small-signal parameters of the MOSFETs, we obtain[14]

[14]Recall that G_{m1} and G_{m2} are the transconductances of, respectively, the first and second stages of the op amp. Capacitors C_1 and C_2 represent the total capacitance to ground at the output nodes of, respectively, the first and second stages of the op amp.

Example PS.13.1 continued

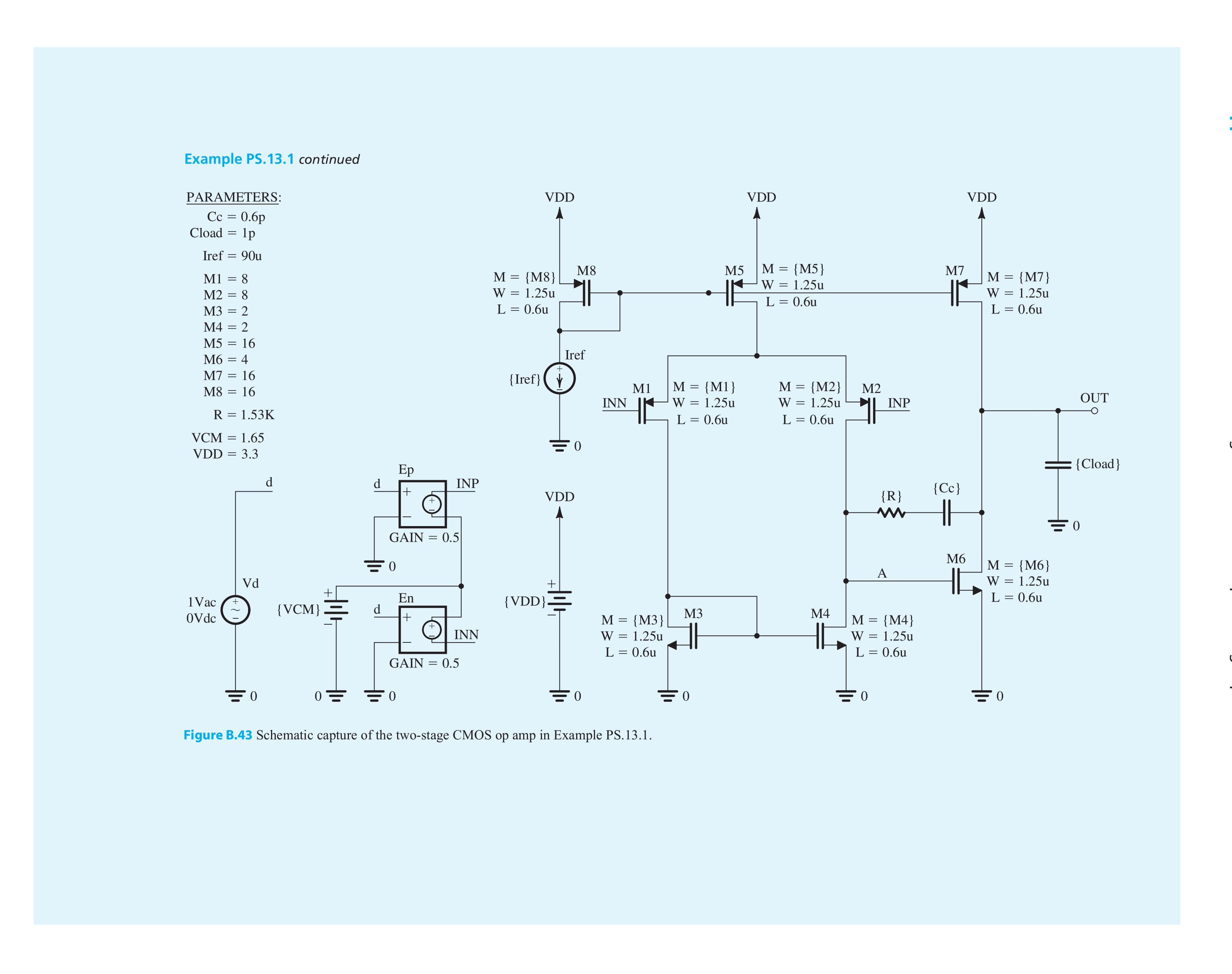


Figure B.43 Schematic capture of the two-stage CMOS op amp in Example PS.13.1.

$$G_{m1} = 0.333\,\text{mA/V}$$
$$G_{m2} = 0.650\,\text{mA/V}$$
$$C_1 = 26.5\,\text{fF}$$
$$C_2 = 1.04\,\text{pF}$$

using Eqs. (13.7), (13.14), (13.25), and (13.26), respectively. Then, using Eq. (13.35), the frequency of the second, nondominant, pole can be found as

$$f_{P2} \simeq \frac{G_{m2}}{2\pi C_2} = 97.2\,\text{MHz}$$

In order to place the transmission zero, given by Eq. (13.31), at infinite frequency, we select

$$R = \frac{1}{G_{m2}} = 1.53\,\text{k}\Omega$$

Now, using Eq. (13.31), the phase margin of the op amp can be expressed as

$$PM = 90° - \tan^{-1}\left(\frac{f_t}{f_{P2}}\right) \tag{B.44}$$

where f_t is the unity-gain frequency, given in Eq. (13.83),

$$f_t = \frac{G_{m1}}{2\pi C_C} \tag{B.45}$$

Using Eqs. (B.44) and (B.45), we determine that compensation capacitors of $C_C = 0.78$ pF and $C_C = 2$ pF are required to achieve phase margins of $PM = 55°$ and $PM = 75°$, respectively.

Next, an ac-analysis simulation is performed in PSpice to compute the frequency response of the op amp and to verify the foregoing design values. It was found that with $R = 1.53$ kΩ, we needed $C_C = 0.6$ pF and $C_C = 1.8$ pF to set $PM = 55°$ and $PM = 75°$, respectively. We note that these values are reasonably close to those predicted by hand analysis. The corresponding frequency responses for the compensated op amp are plotted in Figs. B.44 and B.45. For comparison, we also show the frequency response of the uncompensated op amp ($C_C = 0$). Observe that the unity gain frequency f_t drops from 70.2 MHz to 26.4 MHz as C_C is increased to improve PM (as anticipated from Eq. B.45).

Rather than increasing the compensation capacitor C_C, the value of the series resistor R can be increased to improve the phase margin PM: For a given C_C, increasing R above $1/G_{m2}$ places the transmission zero at a negative real-axis location (Eq. 13.31), where the phase it introduces *adds* to the phase margin. Thus, PM can be improved without affecting f_t. To verify this point, we set C_C to 0.6 pF and simulate the op-amp circuit in PSpice for the cases of $R = 1.53$ kΩ and $R = 3.2$ kΩ. The corresponding frequency response is plotted in Fig. B.46. Observe how f_t is approximately independent of R. However, by increasing R, PM is improved from 55° to 75°.

Increasing the PM is desirable because it reduces the overshoot in the step response of the op amp. To verify this point, we simulate in PSpice the step response of the op amp for $PM = 55°$ and $PM = 75°$. To do that, we connect the op amp in a unity-gain configuration, apply a small (10-mV) pulse signal at the input with very short (1-ps) rise and fall times to emulate a step input, perform a transient-analysis simulation, and plot the output voltage as shown in Fig. B.47. Observe that the overshoot in the step response drops from 15% to 1.4% when the phase margin is increased from 55° to 75°.

We conclude this example by computing SR, the slew rate of the op amp. From Eq. (13.45),

$$SR = 2\pi f_t V_{OV} = \frac{G_{m1}}{C_C} V_{OV} = 166.5\,\text{V}/\mu\text{s}$$

Example PS.13.1 *continued*

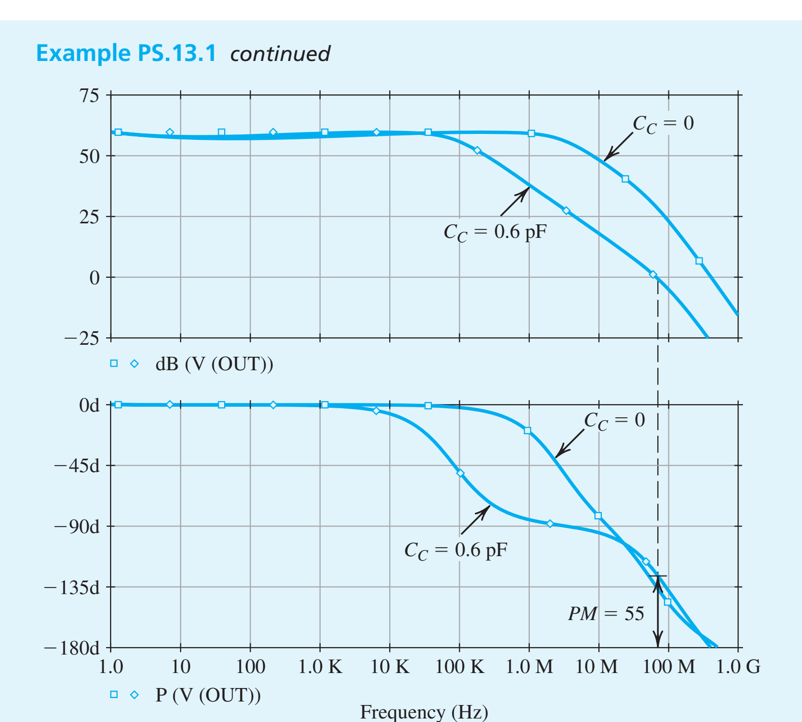


Figure B.44 Magnitude and phase response of the op-amp circuit in Fig. B.43: $R = 1.53\ \text{k}\Omega$, $C_C = 0$ (no frequency compensation), and $C_C = 0.6$ pF ($PM = 55°$).

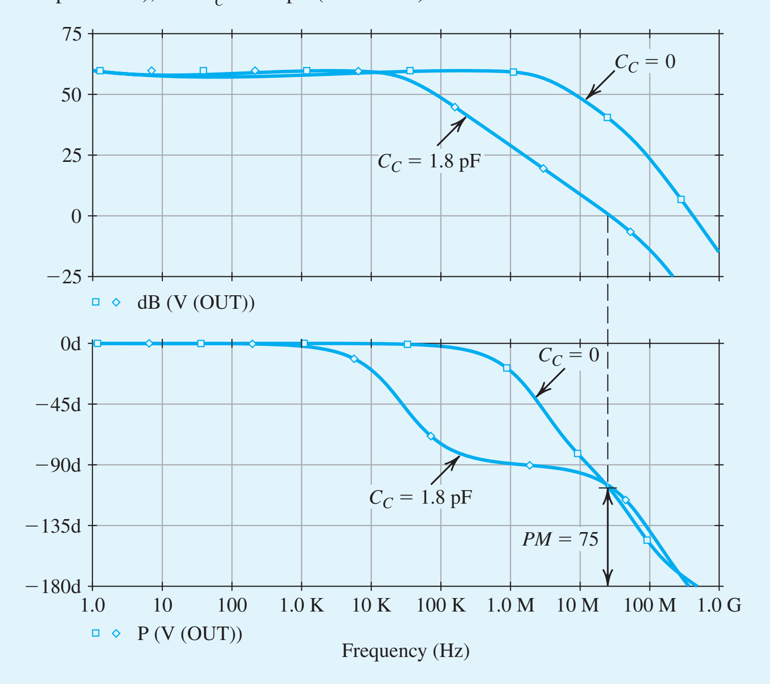


Figure B.45 Magnitude and phase response of the op-amp circuit in Fig. B.43: $R = 1.53\ \text{k}\Omega$, $C_C = 0$ (no frequency compensation), and $C_C = 1.8$ pF ($PM = 75°$).

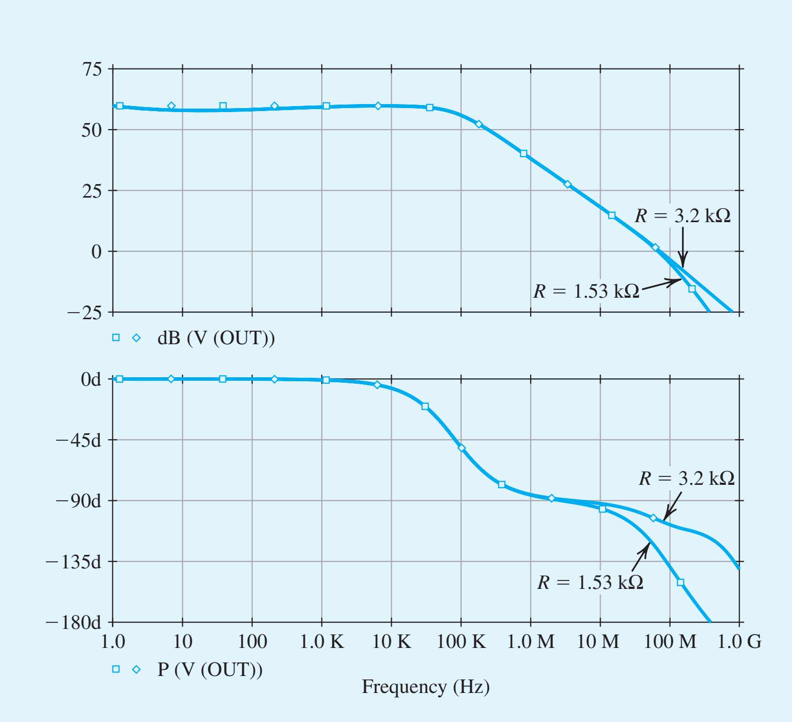


Figure B.46 Magnitude and phase response of the op-amp circuit in Fig. B.43: $C_C = 0.6$ pF, $R = 1.53$ kΩ ($PM = 55°$), and $R = 3.2$ kΩ ($PM = 75°$).

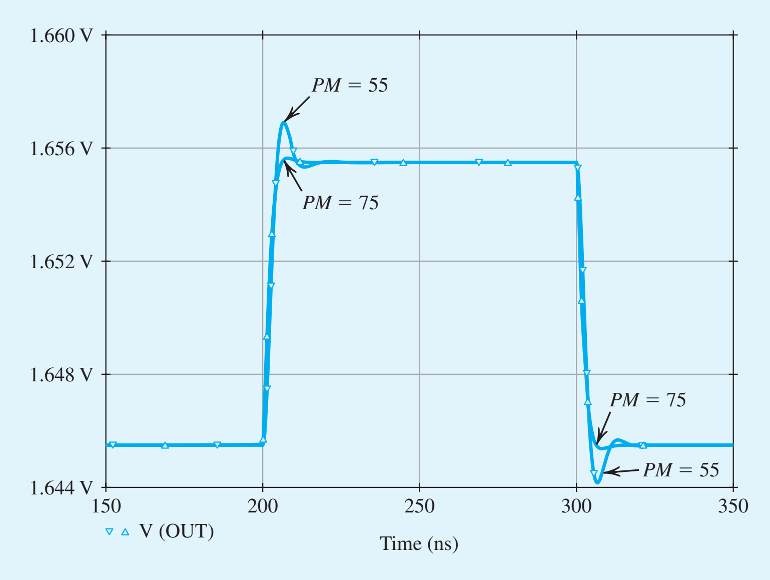


Figure B.47 Small-signal step response (for a 10-mV step input) of the op-amp circuit in Fig. B.43 connected in a unity-gain configuration: $PM = 55°$ ($C_C = 0.6$ pF, $R = 1.53$ kΩ) and $PM = 75°$ ($C_C = 0.6$ pF, $R = 3.2$ kΩ).

Example PS.13.1 *continued*

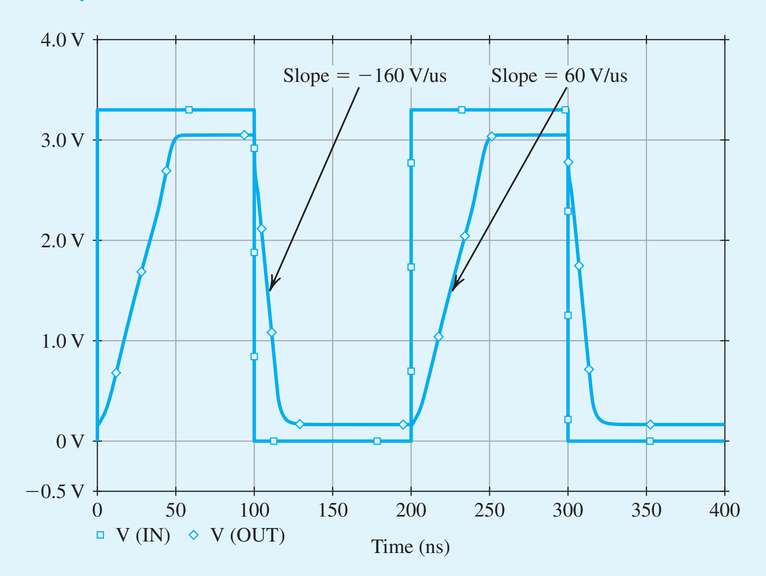


Figure B.48 Large-signal step response (for a 3.3-V step-input) of the op-amp circuit in Fig. B.43 connected in a unity-gain configuration. The slope of the rising and falling edges of the output waveform correspond to the slew rate of the op amp.

when $C_C = 0.6$ pF. Next, to determine *SR* using PSpice (see Example PS.2.2), we again connect the op amp in a unity-gain configuration and perform a transient-analysis simulation. However, we now apply a large pulse signal (3.3 V) at the input to cause slew-rate limiting at the output. The corresponding output-voltage waveform is plotted in Fig. B.48. The slope of the slew-rate-limited output waveform corresponds to the slew rate of the op amp and is found to be $SR = 160$ V/ms and 60 V/ms for the negative- and positive-going output, respectively. These results, with the unequal values of *SR* in the two directions, differ from those predicted by the simple model for the slew-rate limiting of the two-stage op-amp circuit (Section 13.1.6). The difference can perhaps be said to be a result of transistor M_4 entering the triode region and its output current (which is sourced through C_C) being correspondingly reduced. Of course, the availability of PSpice should enable the reader to explore this point further.

Example PS.14.1

Operation of the CMOS Inverter

In this example, we will use PSpice to simulate the CMOS inverter whose schematic capture is shown in Fig, B.49. We will assume a 0.5-µm CMOS technology for the MOSFETs and use parts NMOS0P5 and PMOS0P5 whose level-1 model parameters are listed in Table B.4. In addition to the channel length *L* and the channel width *W*, we have used the multiplicative factor *m* to specify the dimensions of the MOSFETs. The MOSFET parameter *m*, whose default value is 1, is used in SPICE to specify the number of unit-size MOSFETs connected in parallel (see Fig. B.24). In our simulations, we will use unit-size transistors with $L = 0.5$ µm and $W = 1.25$ µm. We will simulate the inverter for two cases: (a) setting $m_p/m_n = 1$ so that the NMOS and PMOS transistors have equal widths, and (b) setting $m_p/m_n = \mu_n/\mu_p = 4$ so that the PMOS

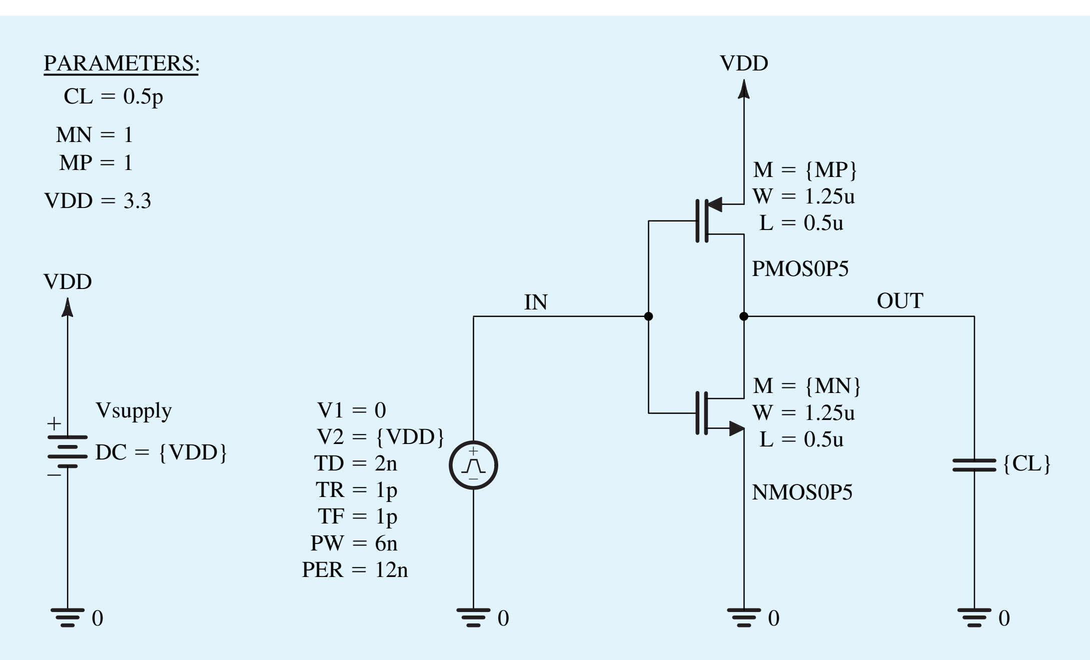


Figure B.49 Schematic capture of the CMOS inverter in Example PS.14.1.

transistor is four times wider than the NMOS transistor (to compensate for the lower mobility in *p*-channel devices as compared with *n*-channel ones). Here, m_n and m_p are the multiplicative factors of, respectively, the NMOS and PMOS transistors of the inverter.

To compute both the voltage transfer characteristic (VTC) of the inverter and its supply current at various values of the input voltage V_{in}, we apply a dc voltage source at the input and perform a dc analysis with V_{in} swept over the range of 0 to V_{DD}. The resulting VTC is plotted in Fig. B.50. Note that the slope of the VTC in the switching region (where both the NMOS and PMOS devices are in saturation) is not infinite as predicted from the simple theory presented in Chapter 14 (Section 14.2, Fig. 14.20). Rather, the nonzero value of λ causes the inverter gain to be finite. Using the derivative feature of Probe, we can find the two points on the VTC at which the inverter gain is unity (i.e., the VTC slope is –1 V/V) and, hence, determine V_{IL} and V_{IH}. Using the results given in Fig. B.50, the corresponding noise margins are $NM_L = NM_H = 1.34$ V for the inverter with $m_p/m_n = 4$, while $NM_L = 0.975$ V and $NM_H = 1.74$ V for the inverter with $m_p/m_n = 1$. Observe that these results correlate reasonably well with the values obtained using the approximate formula in Eq. (14.38). Furthermore, note that with $m_p/m_n = \mu_n/\mu_p = 4$, the NMOS and PMOS devices are closely matched and, hence, the two noise margins are equal.

The threshold voltage V_M of the CMOS inverter is defined as the input voltage v_{IN} that results in an identical output voltage v_{OUT}, that is,

$$V_M = v_{\text{IN}}\big|_{v_{\text{OUT}}=v_{\text{IN}}} \tag{B.46}$$

Thus, as shown in Fig. B.51, V_M is the intersection of the VTC, with the straight line corresponding to $v_{\text{OUT}} = v_{\text{IN}}$ (this line can be simply generated in Probe by plotting v_{IN} versus v_{OUT}, as shown in Fig. B.51). Note that $V_M \simeq (V_{DD}/2)$ for the inverter with $m_p/m_n = 4$. Furthermore, decreasing m_p/m_n decreases V_M. Figure B.51 also shows the inverter supply current versus v_{IN}. Observe that the location of the supply-current peak shifts with the threshold voltage.

Example PS.14.1 *continued*

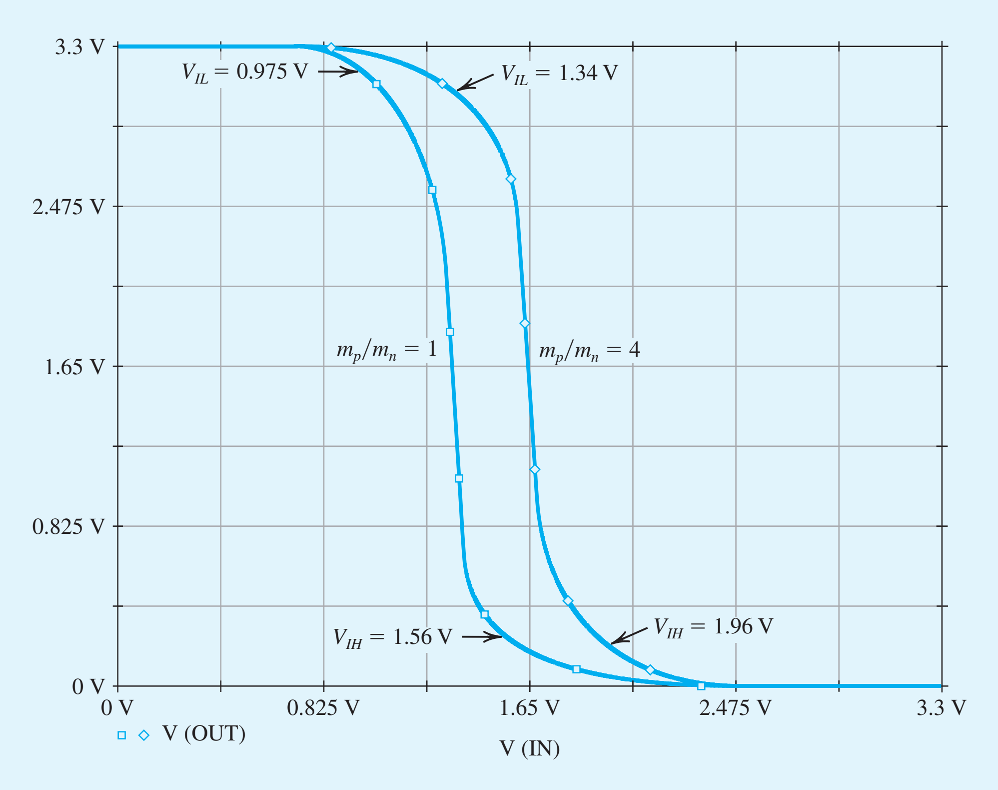
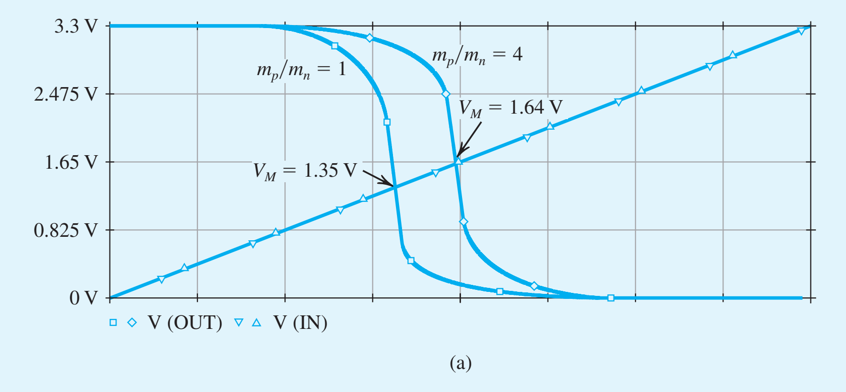
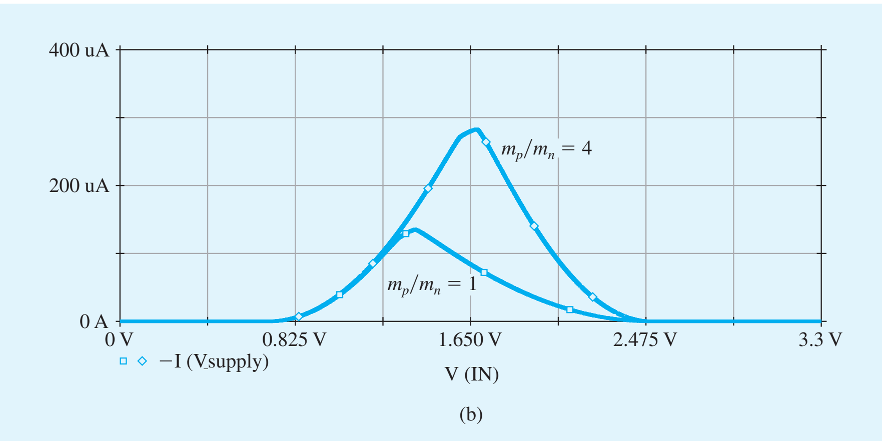


Figure B.50 Input–output voltage transfer characteristic (VTC) of the CMOS inverter in Example PS.14.1 with $m_p/m_n = 1$ and $m_p/m_n = 4$.



(a)

Figure B.51 **(a)** Output voltage and **(b)** supply current versus input voltage for the CMOS inverter in Example PS.14.1 with $m_p/m_n = 1$ and $m_p/m_n = 4$.



(b)

Figure B.51 *continued*

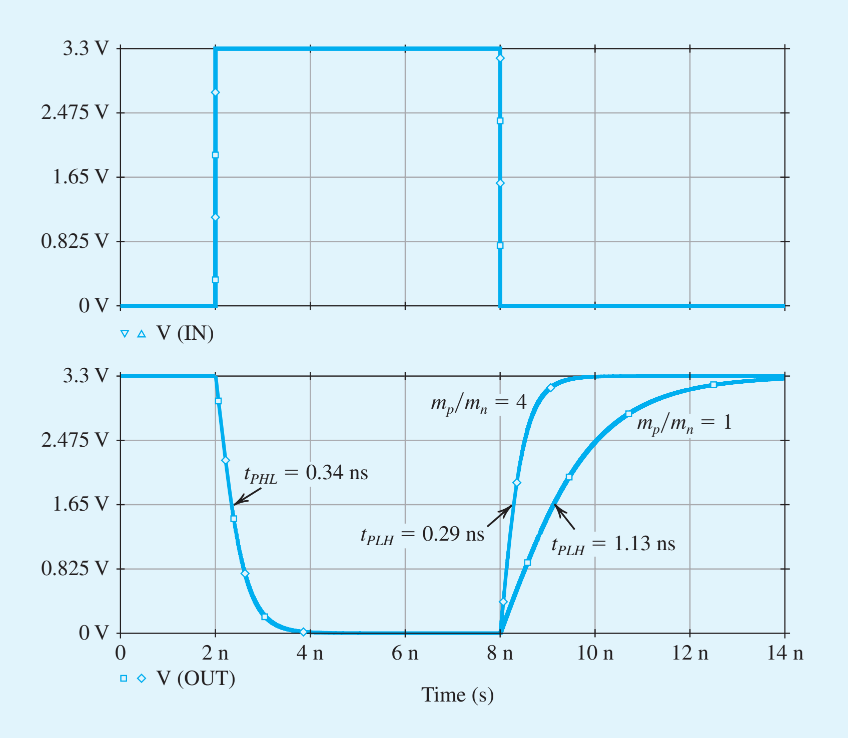


Figure B.52 Transient response of the CMOS inverter in Example PS.13.1 with $m_p/m_n = 1$ and $m_p/m_n = 4$.

To investigate the dynamic operation of the inverter with PSpice, we apply a pulse signal at the input (Fig. B.49), perform a transient analysis, and plot the input and output waveforms as shown in Fig. B.52. The rise and fall times of the pulse source are chosen to be very short. Note that increasing m_p/m_n from 1 to 4 decreases t_{PLH} (from 1.13 ns to 0.29 ns) because of the increased current available to charge C_L, with only a minor increase in t_{PHL} (from 0.33 ns to 0.34 ns). The two propagation delays, t_{PLH} and t_{PHL}, are not exactly equal when $m_p/m_n = 4$, because the NMOS and PMOS transistors are still not perfectly matched (e.g., $V_{tn} \neq |V_{tp}|$).

Example PS.15.1

Static and Dynamic Operation of an ECL Gate

In this example, we use PSpice to investigate the static and dynamic operation of the ECL gate (studied in Section 15.6) whose schematic capture is shown in Fig. B.53.

Having no access to the actual values for the SPICE model parameters of the BJTs utilized in commercially available ECL, we have selected parameter values representative of the technology utilized that, from our experience, would lead to reasonable agreement between simulation results and the measured

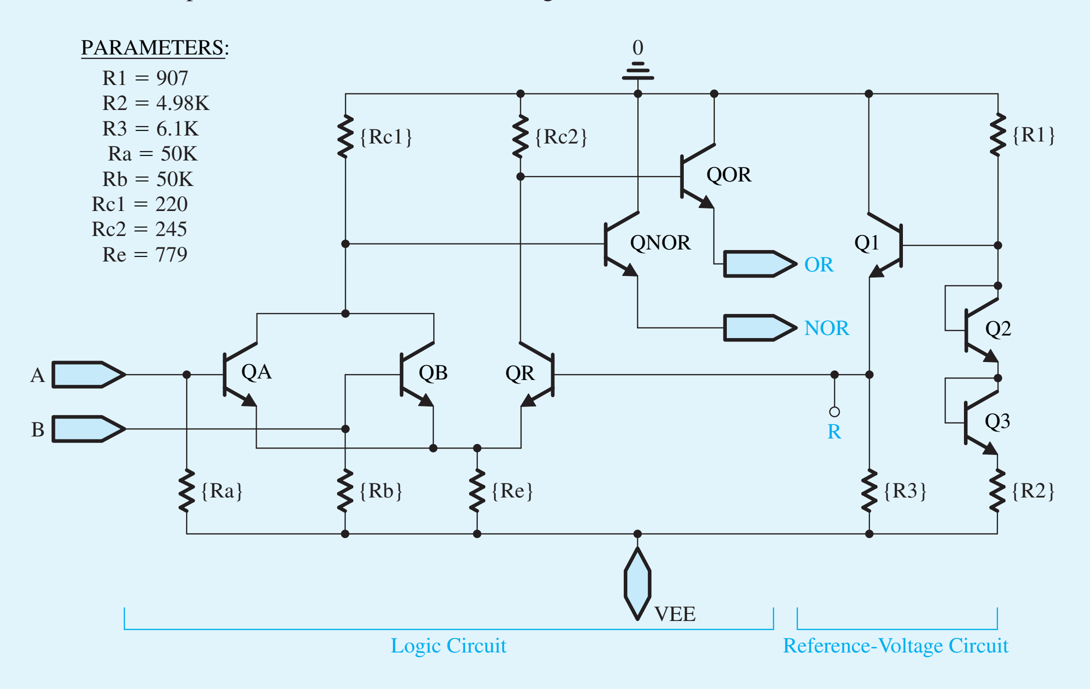


Figure B.53 Schematic capture of the two-input ECL gate for Example PS.15.1

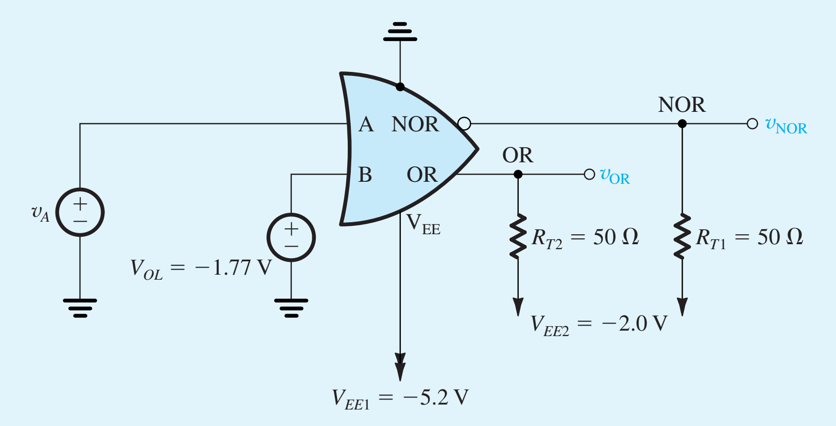


Figure B.54 Circuit arrangement for computing the voltage transfer characteristics of the ECL gate in Fig. B.53.

performance data supplied by the manufacturer. It should be noted that this problem would not be encountered by an IC designer using SPICE as an aid; presumably, the designer would have full access to the proprietary process parameters and the corresponding device model parameters. In any case, for the simulations we conducted, we have utilized the following BJT model parameter values[15]: $I_S = 0.26$ fA, $\beta_F = 100$; $\beta_R = 1$, $\tau_F = 0.1$ ns, $C_{je} = 1$ pF, $C_{jc} = C_\mu = 1.5$ pF, and $|V_A| = 100$ V.

We use the circuit arrangement of Fig. B.54 to compute the voltage transfer characteristics of the ECL gate, that is, v_{OR} and v_{NOR} versus v_A, where v_A is the input voltage at terminal A. For this investigation, the other input is deactivated by applying a voltage $v_B = V_{OL} = -1.77$ V. In PSpice, we perform a dc-analysis simulation with v_A swept over the range –2 V to 0 V in 10-mV increments and plot v_{OR} and v_{NOR} versus v_A. The simulation results are shown in Fig. B.55. We immediately recognize the VTCs as those we have seen and (partially) verified by manual analysis in Section 15.6. The two transfer curves are symmetrical about an input voltage of –1.32 V. PSpice also determined that the voltage V_R at the base of the reference transistor Q_R has exactly this value (–1.32 V), which is also identical to the value we determined by hand analysis of the reference-voltage circuit.

Utilizing Probe (the graphical interface of PSpice), one can determine the values of the important parameters of the VTC, as follows:

OR output: $V_{OL} = -1.77$ V, $V_{OH} = -0.88$ V, $V_{IL} = -1.41$ V, and $V_{IH} = -1.22$ V; thus,
$NM_H = 0.34$ V and $NM_L = 0.36$ V

NOR output: $V_{OL} = -1.78$ V, $V_{OH} = -0.88$ V, $V_{IL} = -1.41$ V, and $V_{IH} = -1.22$ V; thus,
$NM_H = 0.34$ V and $NM_L = 0.37$ V

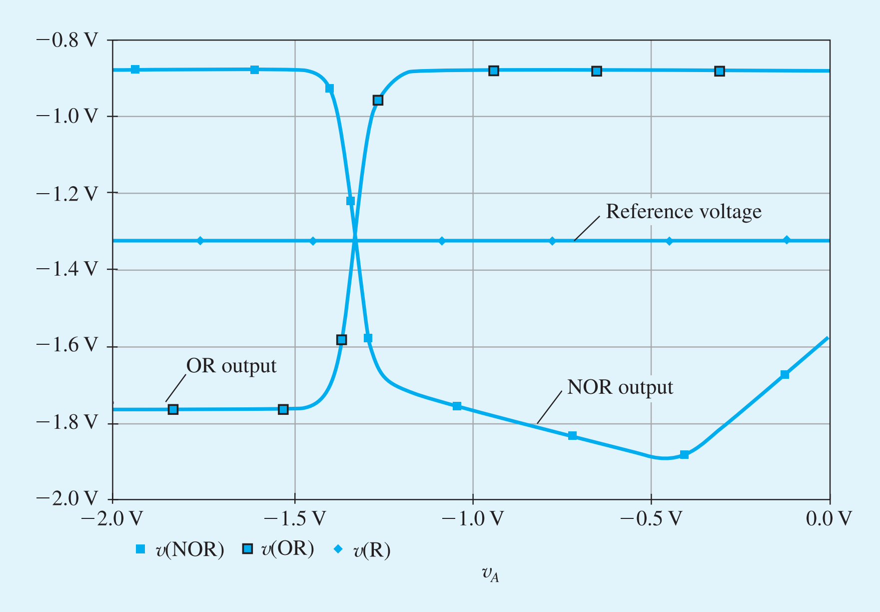


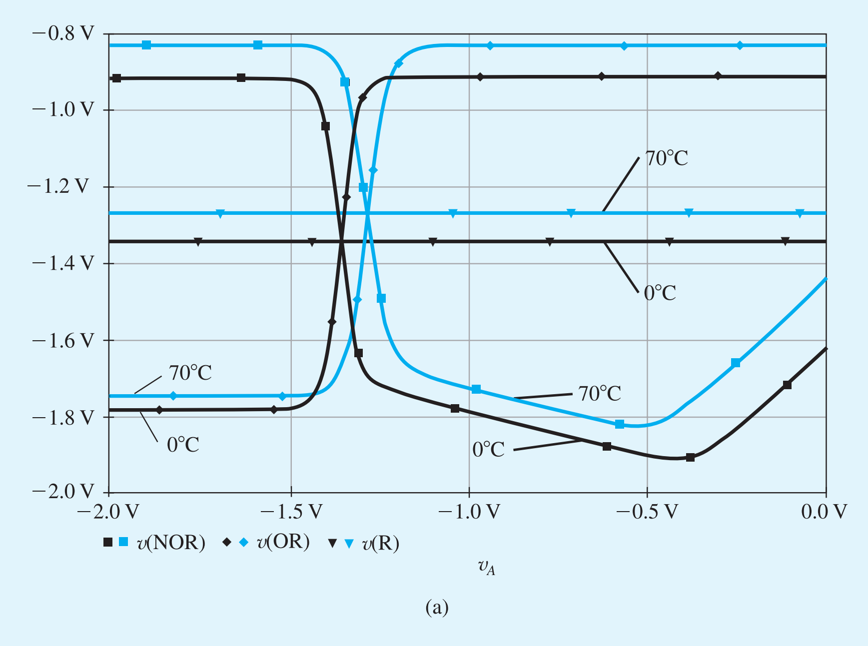
Figure B.55 Voltage transfer characteristics of the OR and NOR outputs (see Fig. B.54) for the ECL gate shown in Fig. B.53. Also indicated is the reference voltage, $V_R = -1.32$ V.

[15]In PSpice, we have created a part called QECL based on these BJT model parameter values. Readers can find this part in the SEDRA.olb library, which is available online at www.oup.com/us/sedrasmith.

Example PS.15.1 *continued*

These values are remarkably close to those found by pencil-and-paper analysis in Section 15.6.

We next use PSpice to investigate the temperature dependence of the transfer characteristics. The reader will recall that in Section 15.6, we discussed this point at some length and carried out a hand analysis in Example 14.4. Here, we use PSpice to find the voltage transfer characteristics at two temperatures, 0°C and 70°C (the VTCs shown in Fig. B.55 were computed at 27°C) for two different cases: the first case with V_R generated as in Fig. B.53, and the second with the reference-voltage circuit eliminated and a constant, temperature-independent reference voltage of –1.32 V applied to the base of Q_R. The simulation results are displayed in Fig. B.56. Figure B.56(a) shows plots of the transfer characteristics for the case in which the reference circuit is utilized, and Fig. B.56(b) shows plots for the case in which a constant reference voltage is employed. Figure B.56(a) indicates that as the temperature is varied and V_R changes, the values of V_{OH} and V_{OL} also change but remain centered on V_R. In other words, the low and high noise margins remain nearly equal. As mentioned in Section 15.6 and demonstrated in the analysis of Example 14.4, this is the basic idea behind making V_R temperature dependent. When V_R is not temperature dependent, the symmetry of V_{OL} and V_{OH} around V_R is no longer maintained, as demonstrated in Fig. B.56(b). Finally, we show some of the values obtained in Table B.12. Observe that for the temperature-compensated case, the average value of V_{OL} and V_{OH} remains very close to V_R. The reader is encouraged to compare these results to those obtained in Example 15.4.



(a)

Figure B.56 Comparing the voltage transfer characteristics of the OR and NOR outputs (see Fig. B.54) of the ECL gate shown in Fig. B.53, with the reference voltage V_R generated using: **(a)** the temperature-compensated bias network of Fig. B.53; **(b)** a temperature-independent voltage source.

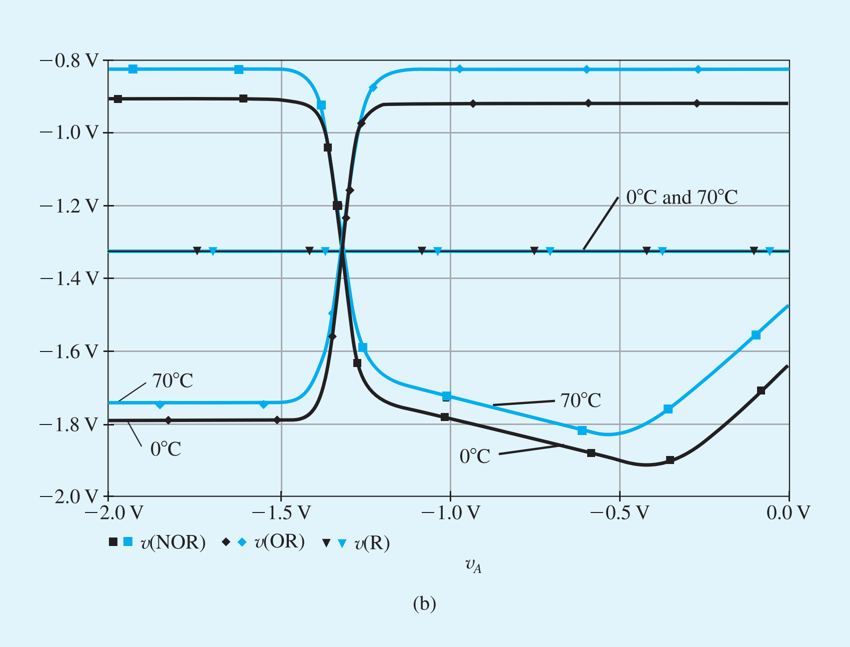


Figure B.56 *continued*

Table B.12 PSpice-Computed Parameter Values of the ECL Gate, with and without Temperature Compensation, at Two Different Temperatures

		Temperature-Compensated		Not Temperature-Compensated	
Temperature	Parameter	OR	NOR	OR	NOR
0°C	V_{OL}	–1.779 V	–1.799 V	–1.786 V	–1.799 V
	V_{OH}	–0.9142 V	–0.9092 V	–0.9142 V	–0.9092 V
	$V_{avg} = \frac{V_{OL} + V_{OH}}{2}$	–1.3466 V	–1.3541 V	–1.3501 V	–1.3541 V
	V_R	–1.345 V	–1.345 V	–1.32 V	–1.32 V
	$\lvert V_{avg} - V_R \rvert$	1.6 mV	9.1 mV	30.1 mV	34.1 mV
70°C	V_{OL}	–1.742 V	–1.759 V	–1.729 V	–1.759 V
	V_{OH}	–0.8338 V	–0.8285 V	–0.8338 V	–0.8285 V
	$V_{avg} = \frac{V_{OL} + V_{OH}}{2}$	–1.288 V	–1.294 V	–1.2814 V	–1.294 V
	V_R	–1.271 V	–1.271 V	–1.32	–1.32 V
	$\lvert V_{avg} - V_R \rvert$	17 mV	23 mV	38 mV	26.2 mV

The dynamic operation of the ECL gate is investigated using the arrangement of Fig. B.57. Here, two gates are connected by a 1.5-m coaxial cable having a characteristic impedance (Z_0) of 50 Ω. The manufacturer specifies that signals propagate along this cable (when it is *properly terminated*) at about half the speed of light, or 15 cm/ns. Thus we would expect the 1.5-m cable we are using to introduce a delay t_d of 10 ns. Observe that in this circuit (Fig. B.57), resistor R_{T1} provides the proper cable termination. The cable is

Example PS.15.1 *continued*

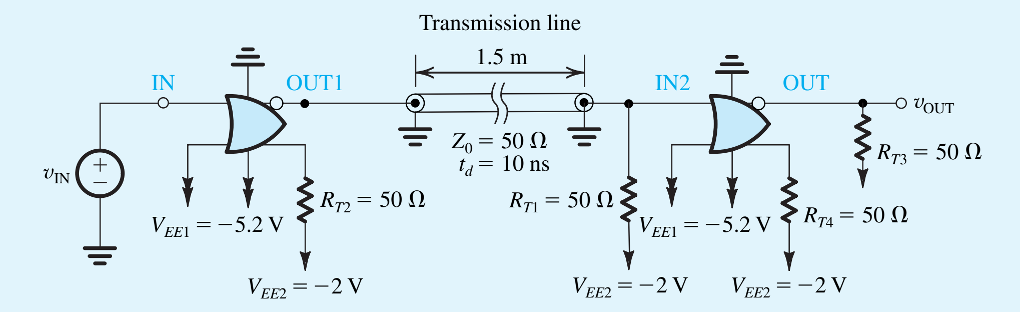


Figure B.57 Circuit arrangement for investigating the dynamic operation of ECL. Two ECL gates (Fig. B.53) are connected in cascade via a 1.5-m coaxial cable, which has a characteristic impedance $Z_0 = 50\ \Omega$ and a propagation delay $t_d = 10$ ns. Resistor R_{T1} (50 Ω) provides proper termination for the coaxial cable.

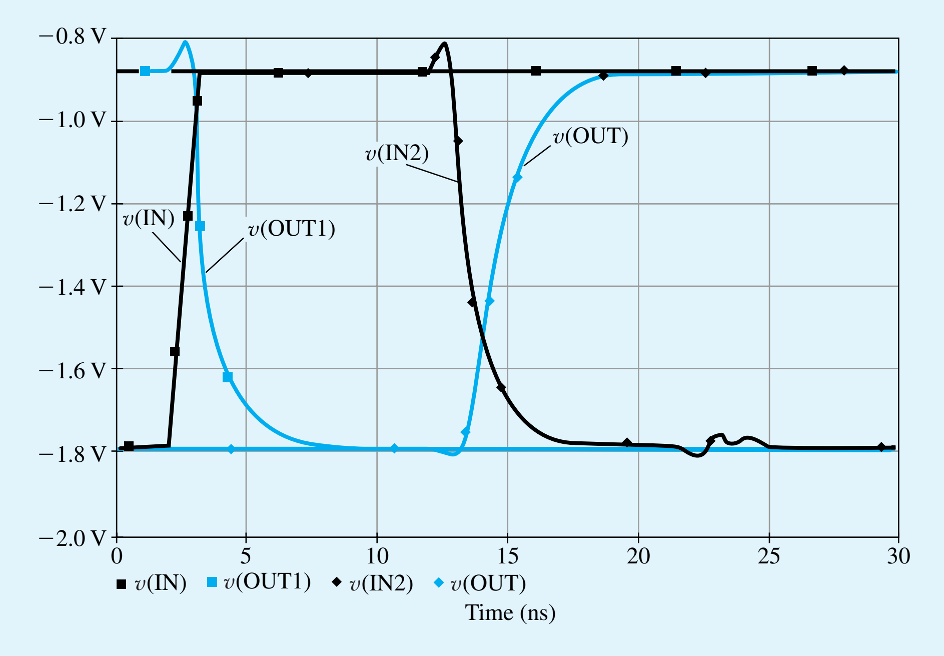


Figure B.58 Transient response of a cascade of two ECL gates interconnected by a 1.5-m coaxial cable having a characteristic impedance of 50 Ω and a delay of 10 ns (see Fig. B.57).

assumed to be lossless and is modeled in PSpice using the *transmission line* element (the T part in the Analog library) with $Z_0 = 50\ \Omega$ and $t_d = 10$ ns. A voltage step, rising from –1.77 V to –0.884 V in 1 ns, is applied to the input of the first gate, and a transient analysis over a 30-ns interval is requested. Figure B.58 shows plots of the waveforms of the input, the voltage at the output of the first gate, the voltage at the input of the second gate, and the output. Observe that despite the very high edge speeds involved, the waveforms are reasonably clean and free of excessive ringing and reflections. This is particularly remarkable because the signal is being transported over a relatively long distance. A detailed examination of the waveforms reveals that the delay along the cable is indeed 10 ns, and the delay of the second gate is about 1.06 ns.

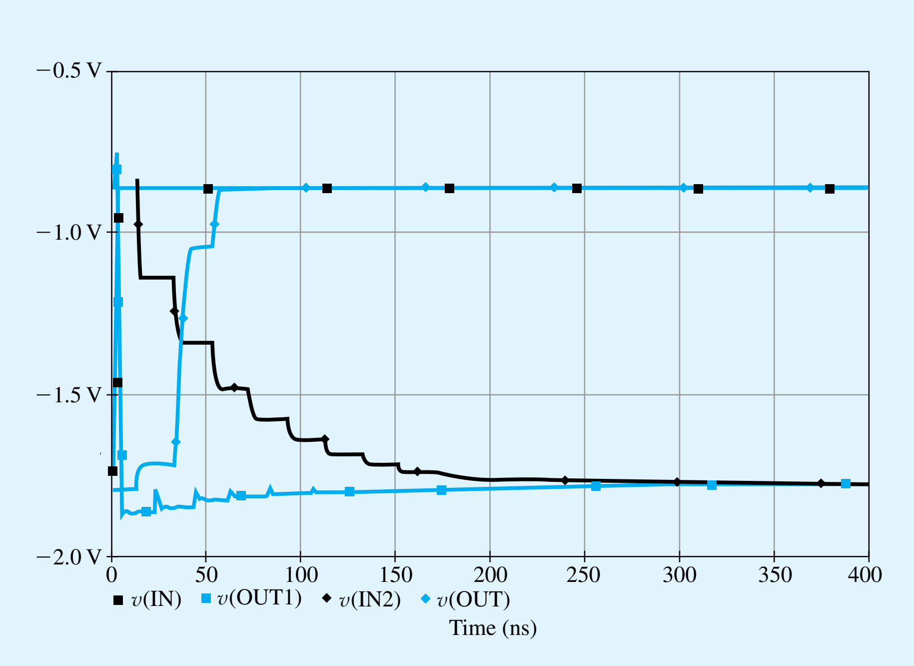


Figure B.59 Transient response of a cascade of two ECL gates interconnected by a 1.5-m cable having a characteristic impedance of 300 Ω. The termination resistance R_{T1} (see Fig. B.57) was kept unchanged at 50 Ω. Note the change in time scale of the plot.

Finally, to verify the need for properly terminating the transmission line, the dynamic analysis is repeated, this time with the 50-Ω coaxial cable replaced with a 300-Ω twisted-pair cable while keeping the termination resistance unchanged. The results are the slow rising and falling and long-delayed waveforms shown in Fig. B.59. (Note the change of plotting scale.)

Example PS.17.1

Verification of the Design of a Fifth-Order Chebyshev Filter

In this example we show how SPICE can be utilized to verify the design of a fifth-order Chebyshev filter. Specifically, we simulate the operation of the circuit whose component values were obtained in Exercise 16.20. The complete circuit is shown in Fig. B.60(a). It consists of a cascade of two second-order simulated-LCR resonators using the Antoniou circuit and a first-order op amp–RC circuit. Using PSpice, we would like to compare the magnitude of the filter response with that computed directly from its transfer function. Here, we note that PSpice can also be used to perform the latter task by using the Laplace transfer-function block in the analog-behavioral-modeling (ABM) library.

Since the purpose of the simulation is simply to verify the design, we assume ideal components. For the op amps, we utilize a near-ideal model, namely, a voltage-controlled voltage source (VCVS) with a gain of 10^6 V/V, as shown in Fig. B.60(b).[16]

[16]SPICE models for the op amp are described in Section B.1.1.

Example PS.17.1 *continued*

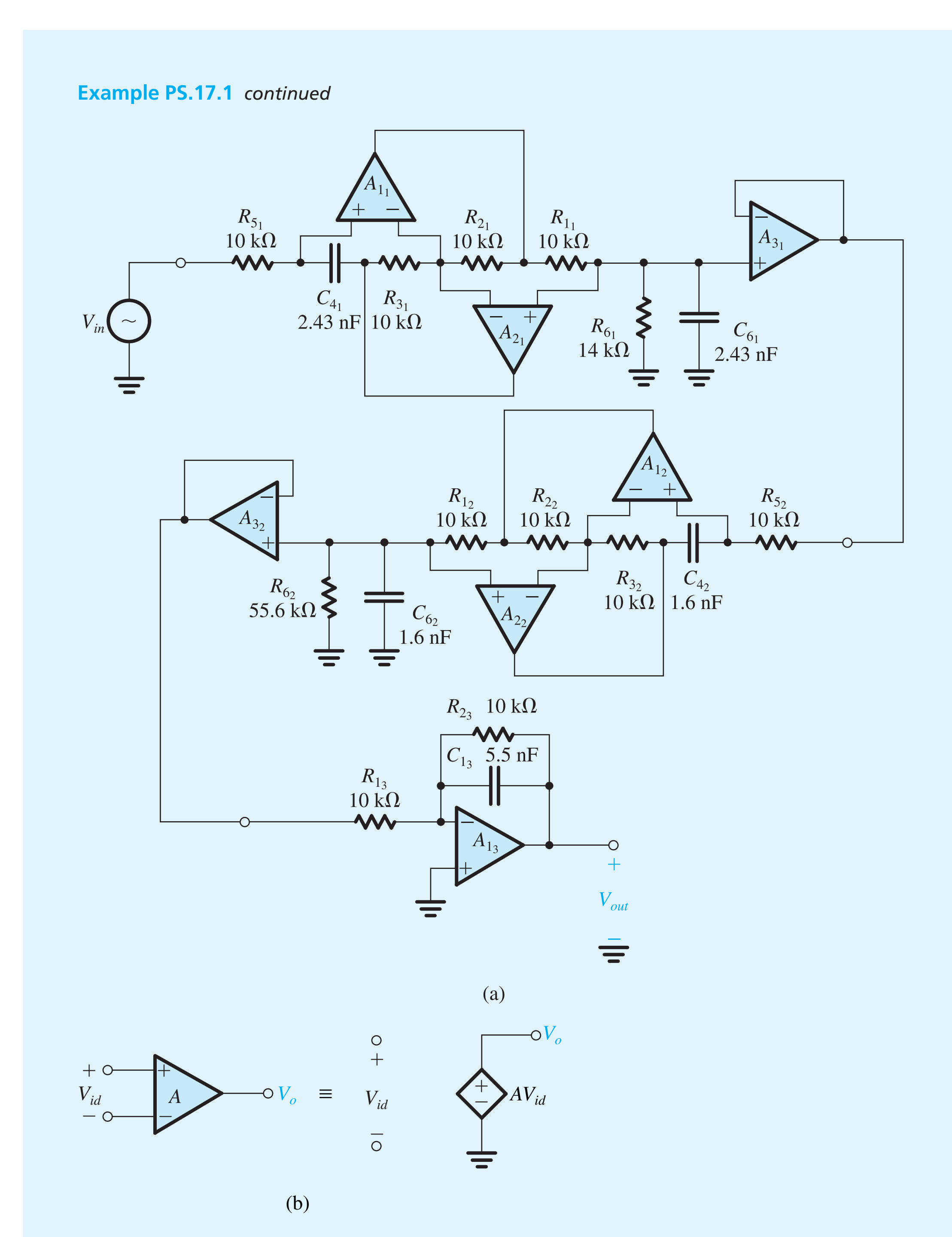


Figure B.60 Circuits for Example PS.17.1. **(a)** Fifth-order Chebyshev filter circuit implemented as a cascade of two second-order simulated LCR resonator circuits and a single first-order op amp–RC circuit. **(b)** VCVS representation of an ideal op amp with gain *A*.

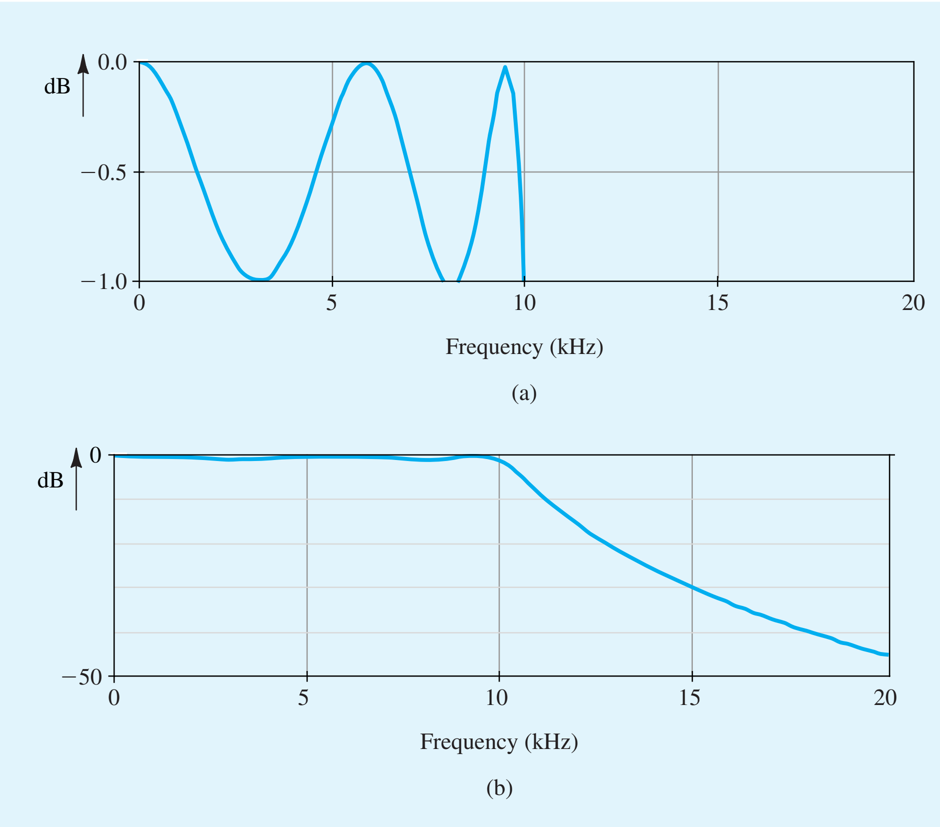


Figure B.61 Magnitude response of the fifth-order lowpass filter circuit shown in Fig. B.60: **(a)** an expanded view of the passband region; **(b)** a view of both the passband and stopband regions.

In SPICE, we apply a 1-V ac signal at the filter input, perform an ac-analysis simulation over the range 1 Hz to 20 kHz, and plot the output voltage magnitude versus frequency, as shown in Fig. B.61. Both an expanded view of the passband and a view of the entire magnitude response are shown. These results are almost identical to those computed directly from the ideal transfer function, thereby verifying the correctness of the design.

Example PS.17.2

Effect of Finite Op-Amp Bandwidth on the Operation of the Two-Integrator-Loop Filter

In this example, we investigate the effect of the finite bandwidth of practical op amps on the response of a two-integrator-loop bandpass filter utilizing the Tow–Thomas biquad circuit of Fig. 16.25(b). The circuit is designed to provide a bandpass response with $f_0 = 10$ kHz, $Q = 20$, and a unity center-frequency gain. The op amps are assumed to be of the 741 type. Specifically, we model the terminal behavior of the op amp with the single-time-constant linear network shown in Fig. B.62. Since the analysis performed here is a small-signal (ac) analysis that ignores nonlinearities, no nonlinearities are included in this op-amp

Example PS.17.2 *continued*

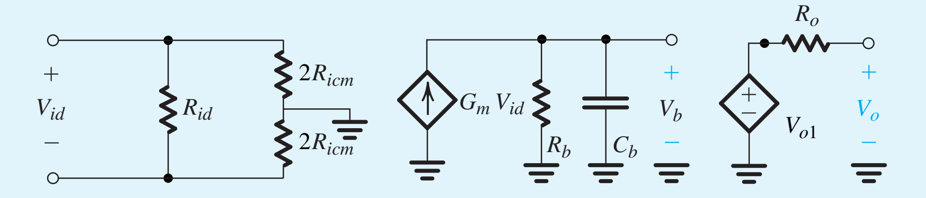


Figure B.62 One-pole equivalent-circuit macromodel of an op amp operated within its linear region.

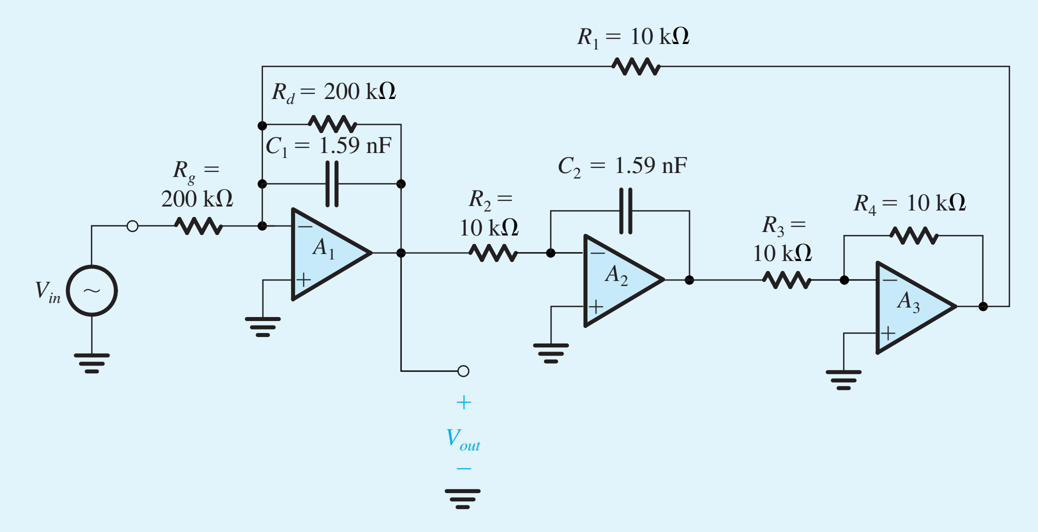


Figure B.63 Circuit for Example PS.17.2 Second-order bandpass filter implemented with a Tow–Thomas biquad circuit having $f_0 = 10$ kHz, $Q = 20$, and unity center-frequency gain.

macromodel. (If the effects of op-amp nonlinearities are to be investigated, a transient analysis should be performed.) The following values are used for the parameters of the op-amp macromodel in Fig. B.62:

$$R_{id} = 2\ M\Omega \qquad R_{icm} = 500\ M\Omega \qquad R_1 = 75\ \Omega$$
$$G_m = 0.19\ mA/V \qquad R_b = 1.323 \times 10^9\ \Omega \qquad C_b = 30\ pF$$

These values result in the specified input and output resistances of the 741-type op amp. Further, they provide a dc gain $A_0 = 2.52 \times 10^5$ V/V and a 3-dB frequency f_b of 4 Hz, again equal to the values specified for the 741. Note that the selection of the individual values of G_m, R_b, and C_b is immaterial as long as $G_m R_b = A_0$ and $C_b R_b = 1/2\pi f_b$.

The Tow-Thomas circuit simulated is shown in Fig. B.6 3. The circuit is simulated in PSpice for two cases: (1) assuming 741-type op amps and using the linear macromodel in Fig. B.62; and (2) assuming ideal op amps with dc gain of $A_0 = 10^6$ V/V and using the near-ideal model in Fig. B.60. In both cases,

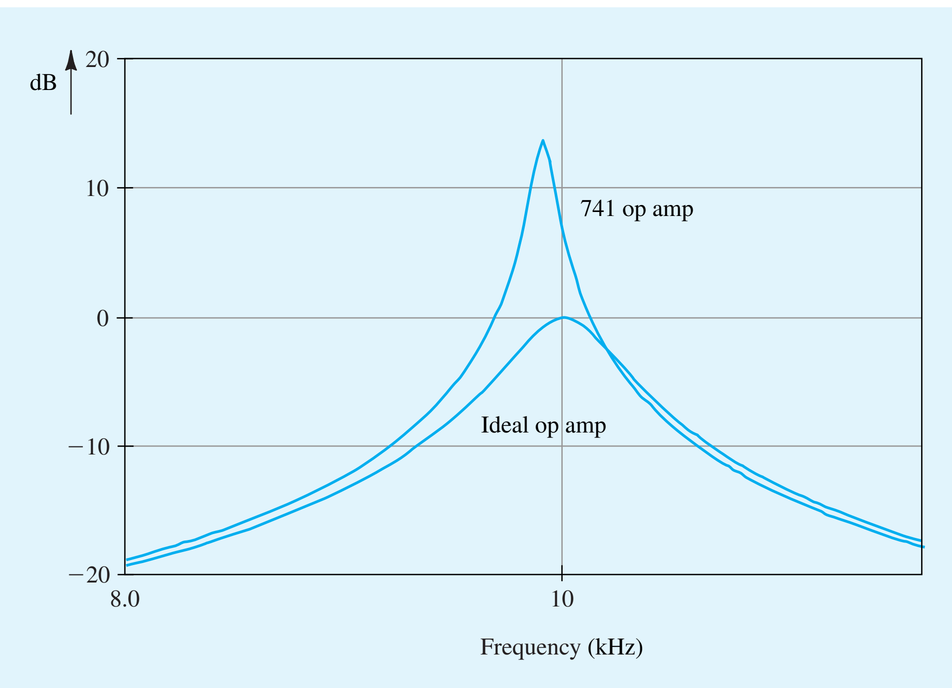


Figure B.64 Comparing the magnitude response of the Tow–Thomas biquad circuit (shown in Fig. B.63) constructed with 741-type op amps, with the ideal magnitude response. These results illustrate the effect of the finite dc gain and bandwidth of the 741 op amp on the frequency response of the Tow–Thomas biquad circuit.

we apply a 1-V ac signal at the filter input, perform an ac-analysis simulation over the range 8 kHz to 12 kHz, and plot the output-voltage magnitude versus frequency.

The simulation results are shown in Fig. B.64, from which we observe the significant deviation between the response of the filter using the 741 op amp and that using the near-ideal op-amp model. Specifically, the response with practical op amps shows a deviation in the center frequency of about –100 Hz, and a reduction in the 3-dB bandwidth from 500 Hz to about 110 Hz. Thus, in effect, the filter Q factor has increased from the ideal value of 20 to about 90. This phenomenon, known as *Q-enhancement,* is predictable from an analysis of the two-integrator-loop biquad with the finite op-amp bandwidth taken into account [see Sedra and Brackett (1978)]. Such an analysis shows that Q-enhancement occurs as a result of the excess phase lag introduced by the finite op-amp bandwidth. The theory also shows that the Q-enhancement effect can be compensated for by introducing phase lead around the feedback loop. This can be accomplished by connecting a small capacitor, C_C, across resistor R_2. To investigate the potential of such a compensation technique, we repeat the PSpice simulation with various capacitance values. The results are displayed in Fig. B.65(a). We observe that as the compensation capacitance is increased from 0 pF, both the filter Q and the resonance peak of the filter response move closer to the desired values. It is evident, however, that a compensation capacitance of 80 pF causes the response to deviate further from the ideal. Thus, optimum compensation is obtained with a capacitance value between 60 pF and 80 pF. Further experimentation using PSpice enabled us to determine that such an optimum is obtained with a compensation capacitance of 64 pF. The corresponding response is shown, together with the ideal response, in Fig. B.65(b). We note that although the filter Q has been restored to its ideal value, there remains a deviation in the center frequency. We shall not pursue this matter any further

Example PS.17.2 *continued*

here; our objective is not to present a detailed study of the design of two-integrator-loop biquads; rather, it is to illustrate the application of SPICE in investigating the nonideal performance of active-filter circuits, generally.

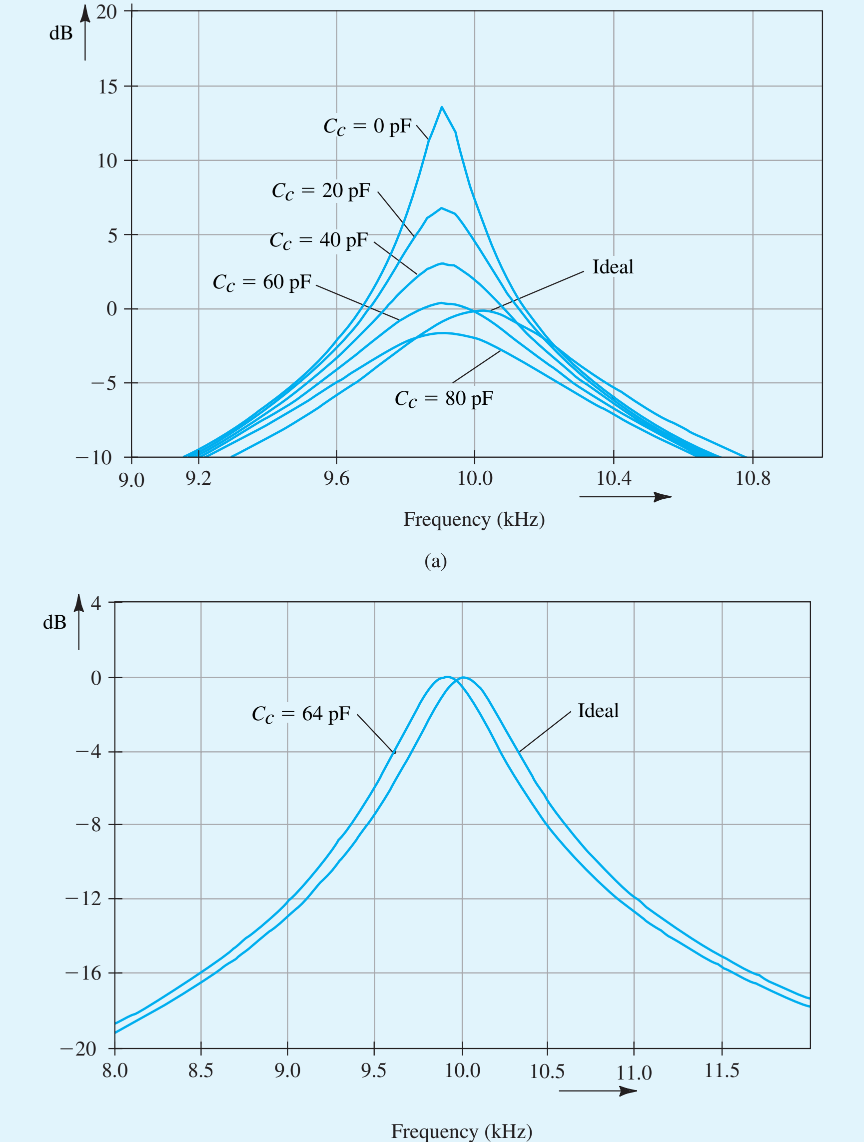


Figure B.65 **(a)** Magnitude response of the Tow–Thomas biquad circuit with different values of compensation capacitance. For comparison, the ideal response is also shown. **(b)** Comparing the magnitude response of the Tow–Thomas biquad circuit using a 64-pF compensation capacitor and the ideal response.

Example PS.18.1

Wien-Bridge Oscillator

For our first example on oscillators, we shall simulate the operation of the Wien-bridge oscillator whose schematic capture is shown in Fig. B.66. The component values are selected to yield oscillations at 1 kHz. We would like to investigate the operation of the circuit for different settings of R_{1a} and R_{1b}, with $R_{1a}+R_{1b}=50\,\text{k}\Omega$. Since oscillation just starts when $(R_2+R_{1b})/R_{1a}=2$ (see Exercise 17.4), that is, when $R_{1a}=20$ kΩ and $R_{1b}=30$ kΩ, we consider three possible settings: (a) $R_{1a}=15$ kΩ, $R_{1b}=35$ kΩ; (b) $R_{1a}=18$ kΩ, $R_{1b}=32$ kΩ; and (c) $R_{1a}=25$ kΩ, $R_{1b}=25$ kΩ. These settings correspond to loop gains of 1.33, 1.1, and 0.8, respectively.

In PSpice, a 741-type op amp and 1N4148-type diodes are used to simulate the circuit similar to the one in Fig. 18.7.[17] A transient-analysis simulation is performed with the capacitor voltages initially set to zero. This demonstrates that the op-amp offset voltage is sufficient to cause the oscillations to start without the need for special start-up circuitry. Figure B.67 shows the simulation results. The graph in Fig. B.67(a)

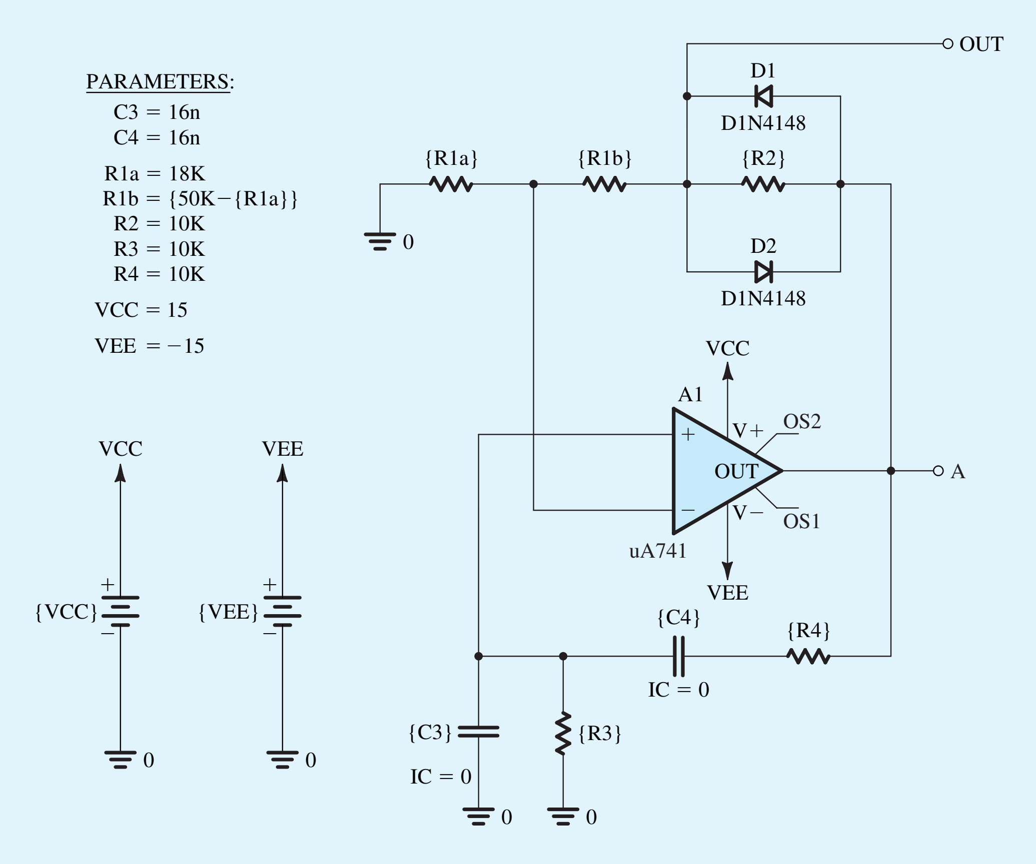
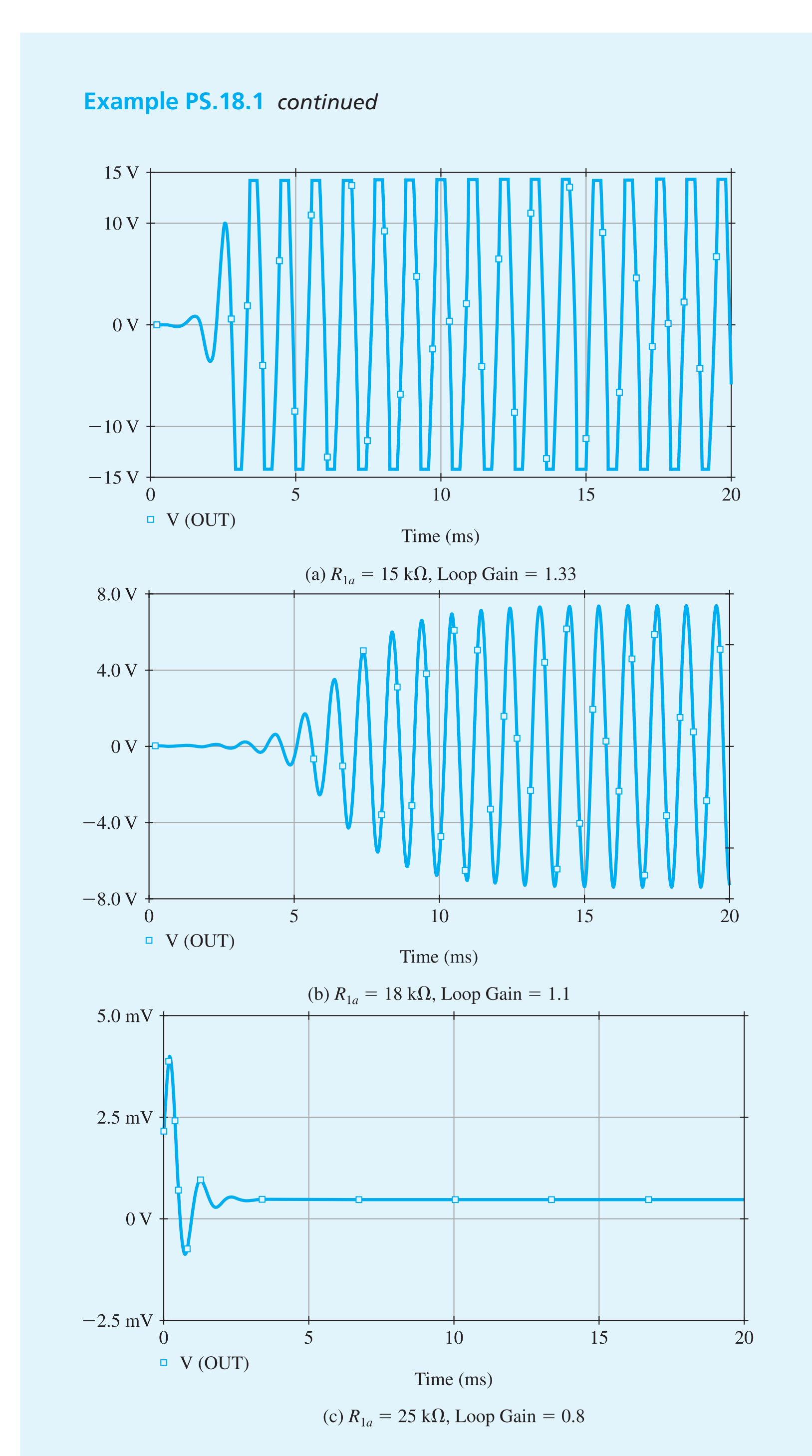


Figure B.66 Example PS.18.1: Schematic capture of a Wien-bridge oscillator.

[17]The SPICE models for the 741 op amp and the 1N4148 diode are available in PSpice. The 741 op amp was characterized in Example PS.2.2. The 1N4148 diode was used in Example PS.4.1.

Example PS.18.1 *continued*



(a) R_{1a} = 15 kΩ, Loop Gain = 1.33

(b) R_{1a} = 18 kΩ, Loop Gain = 1.1

(c) R_{1a} = 25 kΩ, Loop Gain = 0.8

Figure B.67 Start-up transient behavior of the Wien-bridge oscillator shown in Fig. B.66 for various values of loop gain.

shows the output waveform obtained for a loop gain of 1.33 V/V. Observe that although the oscillations grow and stabilize rapidly, the distortion is considerable. The output obtained for a loop gain of 1.1, shown in Fig. B.67(b), is much less distorted. However, as expected, as the loop gain is reduced toward unity, it takes longer for the oscillations to build up and for the amplitude to stabilize. For this case, the frequency is 986.6 Hz, which is reasonably close to the design value of 1 kHz, and the amplitude is 7.37 V. Finally, for a loop gain of 0.8, the output shown in Fig. B.67(c) confirms our expectation that sustained oscillations cannot be obtained when the loop gain is less than unity.

PSpice can be used to investigate the spectral purity of the output sine wave. This is achieved using the Fourier analysis facility. It is found that in the steady state, the output for the case of a loop gain of 1.1 has a THD figure of 1.88%. When the oscillator output is taken at the op-amp output (voltage v_A), a THD of 2.57% is obtained, which, as expected, is higher than that for the voltage v_{OUT}, but not by very much. The output terminal of the op amp is of course a much more convenient place to take the output.

Example PS.18.2

Active-Filter-Tuned Oscillator

In this example, we use PSpice to verify our contention that a superior op amp–oscillator can be realized using the active-filter-tuned circuit of Fig. 18.11. We also investigate the effect of changing the value of the filter Q factor on the spectral purity of the output sine wave.

Consider the circuit whose schematic capture is shown in Fig. B.68. For this circuit, the center frequency is 1 kHz, and the filter Q is 5 when $R_1 = 50$ kΩ and 20 when $R_1 = 200$ kΩ. As in the case of the Wien-bridge circuit in Example PS.18.1, 741-type op amps and 1N4148-type diodes are utilized. In PSpice, a transient-analysis simulation is performed with the capacitor voltages initially set to zero. To be able to compute the Fourier components of the output, the analysis interval chosen must be long enough to allow the oscillator to reach a steady state. The time to reach a steady state is in turn determined by the value of the filter Q; the higher the Q, the longer it takes the output to settle. For $Q = 5$, it was determined, through a combination of approximate calculations and experimentation using PSpice, that 50 ms is a reasonable estimate for the analysis interval. For plotting purposes, we use 200 points per period of oscillation.

The results of the transient analysis are plotted in Fig. B.69. The upper graph shows the sinusoidal waveform at the output of op amp A_1 (voltage v_1). The lower graph shows the waveform across the diode limiter (voltage v_2). The frequency of oscillation is found to be very close to the design value of 1 kHz. The amplitude of the sine wave is determined using Probe (the graphical interface of PSpice) to be 1.15 V (or 2.3 V p-p). Note that this is lower than the 3.6 V estimated in Exercise 17.7. The latter value, however, was based on an estimate of 0.7-V drop across each conducting diode in the limiter. The lower waveform in Fig. B.69 indicates that the diode drop is closer to 0.5 V for a 1 V peak-to-peak amplitude of the pseudo-square wave. We should therefore expect the peak-to-peak amplitude of the output sinusoid to be lower than 3.6 V by the same factor, and indeed it is approximately the case.

In PSpice, the Fourier analysis of the output sine wave indicates that THD = 1.61%. Repeating the simulation with Q increased to 20 (by increasing R_1 to 200 kΩ), we find that the value of THD is reduced to 1.01%. Thus, our expectations that the value of the filter Q can be used as an effective means for controlling the THD of the output waveform are confirmed.

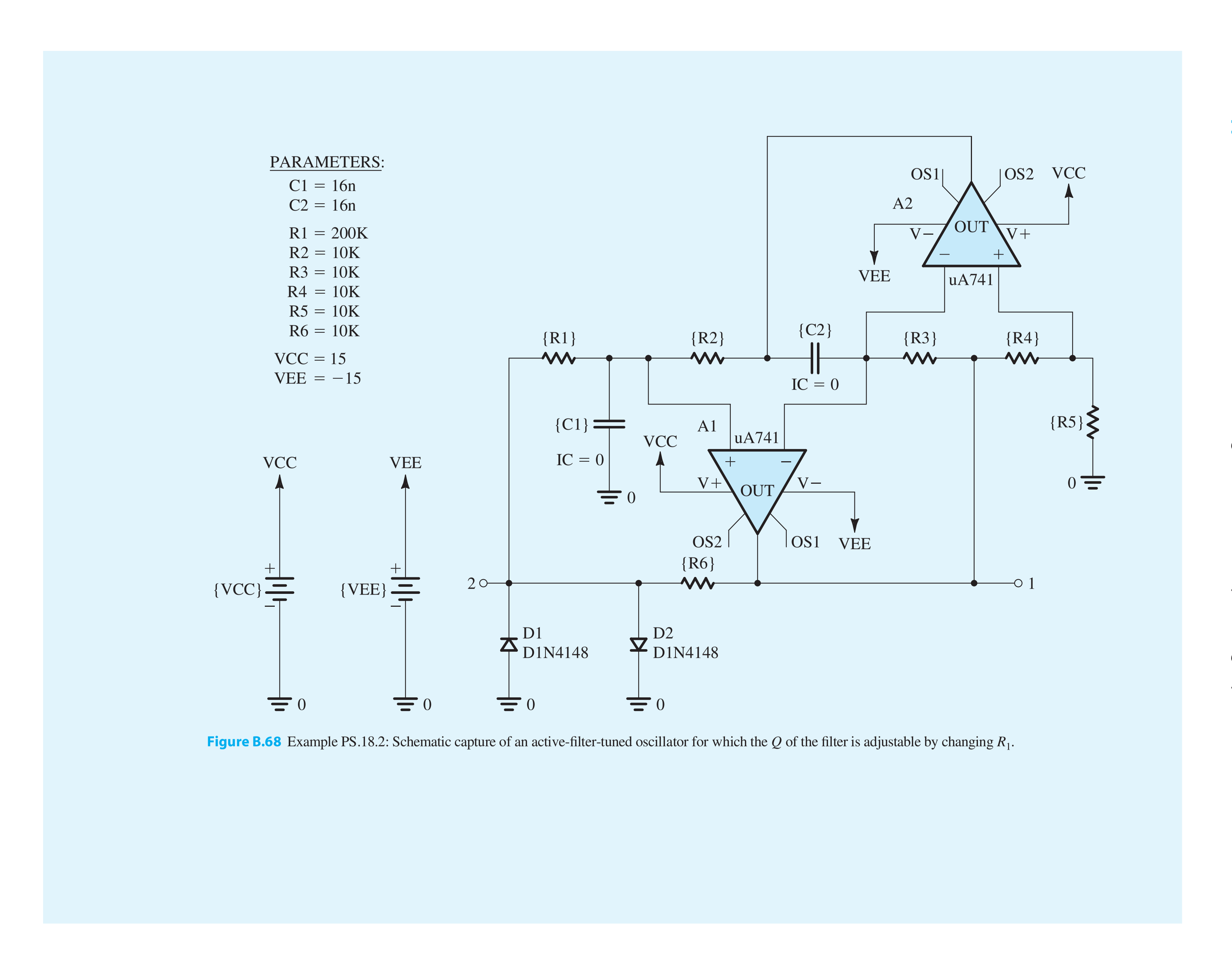


Figure B.68 Example PS.18.2: Schematic capture of an active-filter-tuned oscillator for which the Q of the filter is adjustable by changing R_1.

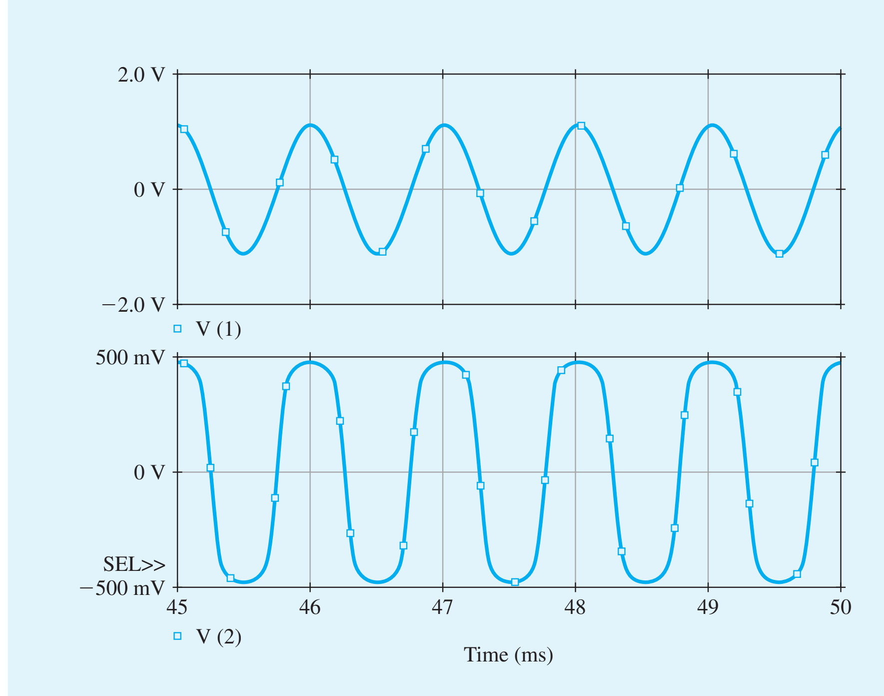


Figure B.69 Output waveforms of the active-filter-tuned oscillator shown in Fig. B.68 for $Q = 5$ ($R_1 = 50$ kΩ).

B.3 Multisim Examples

Example MS.7.1

The CS Amplifier

In this example, we will use Multisim to characterize a CS amplifier whose schematic capture is shown in Fig. B.70. We will assume a 0.18-μm CMOS technology for the MOSFET and use typical SPICE level-1 model parameters for this technology, as provided in Table B.4. We will also assume a signal-source resistance $R_{\text{sig}} = 10$ kΩ, a load resistance $R_L = 50$ kΩ, and bypass and coupling capacitors of 10 μF. The targeted specifications for this CS amplifier are a voltage gain $|A_v| = 10$ V/V and a maximum power consumption $P = 0.45$ mW. As should always be the case with computer simulation, we will begin with an approximate hand-analysis design. We will then use Multisim to fine-tune our design and to investigate the performance of the final design.

The amplifier specifications are summarized in Table B.13.

Hand Design

With a 1.8-V power supply, the drain current of the MOSFET must be limited to $I_D = P/V_{DD} = 0.45\,\text{mW}/1.8\text{ V} = 0.25\,\text{mA}$ to meet the power consumption specification. Choosing $V_{OV} = 0.15$ V and $V_{DS} = V_{DD}/3 = 0.6$ V (to achieve a large signal swing at the output), the MOSFET can now be sized as

Example MS.7.1 *continued*

DEVICE PARAMETERS	
NAME	Q1:NMOS
W	15.48u
L	0.2u
KP	291u
LD	0.01u
VID	0.45
LAMBDA	0.08
GAMMA	0.3

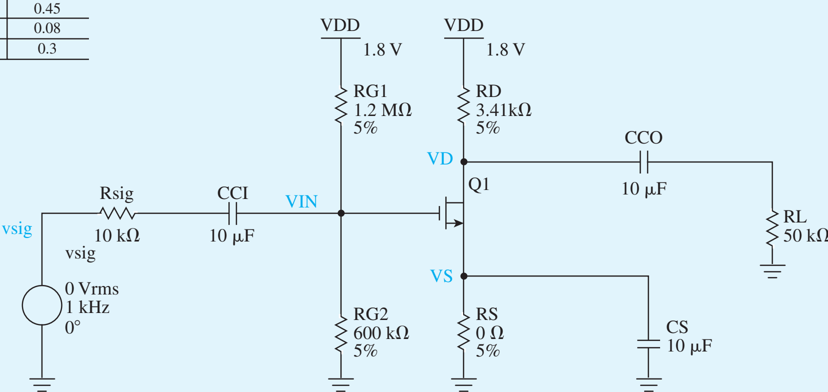


Figure B.70 Capture schematic of the CS amplifier.

Table B.13 CS Amplifier Specifications

Parameters	Value
Power	0.45 mW
R_{sig}	10 kΩ
R_L	50 kΩ
$\lvert A_v \rvert$	10 V/V
V_{DD}	1.8 V

$$\frac{W}{L_{\text{eff}}} = \frac{I_D}{\frac{1}{2}k_n' V_{OV}^2 (1+\lambda V_{DS})} = \frac{250 \times 10^{-6}}{\frac{1}{2} \times 246.2 \times 10^{-2} \times 0.15^6 \times (1+0.08 \times 0.6)} \simeq 86$$

where $k_n' = \mu_n C_{\text{ox}} = 246.2\,\mu\text{A/V}^2$. Here, L_{eff} rather than L is used to more accurately compute I_D.

The effect of using W_{eff} instead of W is much less important, because typically $W \gg W_{ov}$. Thus, choosing $L = 0.200\,\mu\text{m}$ results in $L_{\text{eff}} = L - 2L_{ov} = 0.180\,\mu\text{m}$, and $W = 86 \times L_{\text{eff}} = 15.48\,\mu\text{m}$.

Note that we chose L slightly larger than $L_{\min}$. This is a common practice in the design of analog ICs to minimize the effects of fabrication nonidealities on the actual value of L. As we have seen, this is particularly important when the circuit performance depends on the matching between the dimensions of two or more MOSFETs (e.g., in the current-mirror circuits studied in Chapter 8).

Next, R_D is calculated based on the desired voltage gain:

$$|A_v| = g_m (R_D \| R_L \| r_o) = 10\text{V/V} \Rightarrow R_D \simeq 3.41\text{ k}\Omega$$

where

$$g_m = \frac{2I_D}{V_{OV}} = \frac{2 \times 0.25 \times 10^{-3}}{0.15} = 3.33\,\text{mA/V}$$

and

$$r_o = \frac{V_A}{I_D} = \frac{12.5}{0.25 \times 10^{-3}} = 50\,\text{k}\Omega$$

Hence, the dc bias voltage is $V_D = V_{DD} - I_D R_D = 0.9457\,\text{V}$.

To stabilize the bias point of the CS amplifier, we include a resistor in the source lead. In other words, to bias the MOSFET at $V_{DS} = V_{DD}/3$, we need an

$$R_s = \frac{V_S}{I_D} = \frac{(V_D - V_{DD}/3)}{I_D} = \frac{0.3475}{0.25 \times 10^{-3}} = 1.39\,\text{k}\Omega$$

However, as a result of including such a resistor, the gain drops by a factor of $(1 + g_m R_S)$. Therefore, we include a capacitor, C_S, to eliminate the effect of R_S on ac operation of the amplifier and gain.

Finally, choosing the current in the biasing branch to be 1 µA gives $R_{G1} + R_{G2} = V_{DD}/1\,\mu\text{A} = 1.8\,\Omega$. Also, we know that

$$V_{GS} = V_{OV} + V_t = 0.15 + 0.45 = 0.6\,\text{V} \Rightarrow V_G = V_S + 0.6 = 0.3475 + 0.6 = 0.9475\,\text{V}$$

Hence,

$$\frac{R_{G2}}{R_{G1} + R_{G2}} = \frac{V_G}{V_{DD}} = \frac{0.9475}{1.8} \Rightarrow R_{G1} = 0.8525\,\text{M}\Omega, R_{G2} = 0.9475\,\text{M}\Omega$$

Using large values for these gate resistors ensures that both their power consumption and the loading effect on the input signal source are negligible.

Simulation

Amplifier Biasing We will now use Multisim to verify our design and investigate the performance of the CS amplifier. We begin by performing a bias-point simulation to verify that the MOSFET is properly biased in the saturation region and that the dc voltages and currents match the expected values (refer to this example's simulation file: Ch5_CS_Amplifier_Ex_DC.ms10). The results are shown in Fig. B.71.

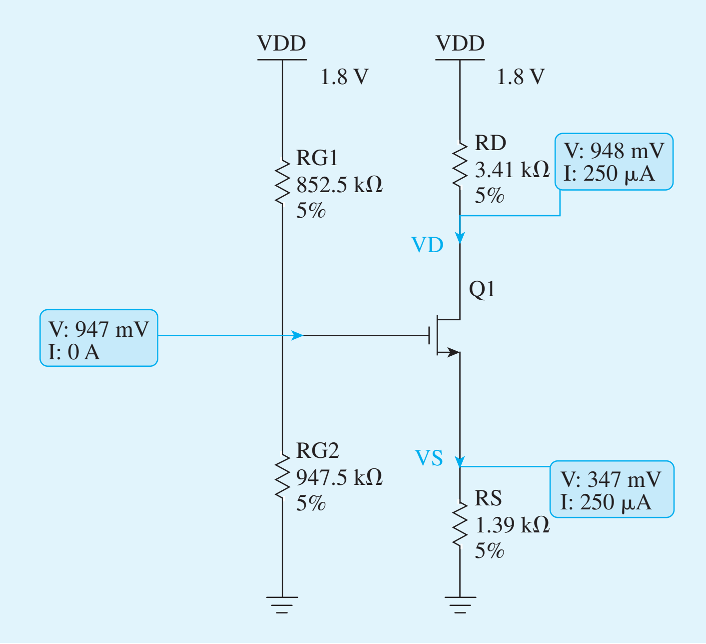


Figure B.71 DC bias-point analysis of the CS amplifier.

Example MS.7.1 *continued*

Amplifier Gain We can also verify if our design provides the desired gain. This can be done by performing transient response analysis, as set up in Ch5_CS_Amplifier_Ex_gain.ms10. As can be seen from Fig. B.72, $|G_v| \simeq |A_v| \simeq 11$ V/V. Note the values of overall voltage gain G_v and A_v are close since $R_{in} = (R_{G1} \| R_{G2}) \gg R_{sig}$. In the case where the capacitor C_S is not included ($C_S = 0$), the gain drops by a factor of 5.63 (approximately equal to $1 + g_m R_S$) to 1.95. This is as expected from our study of the CS amplifier with a source-degeneration resistance.

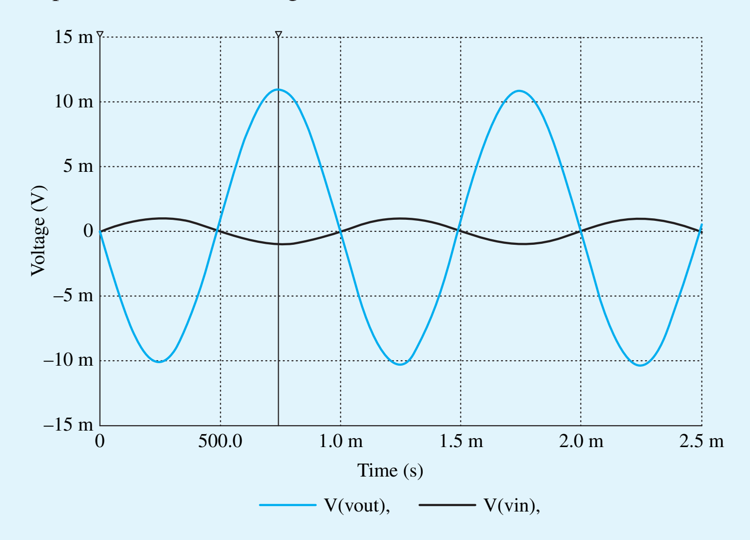


Figure B.72 A_v and G_v of the CS amplifier: transient analysis.

Investigating Amplifier Bias Stability We can also demonstrate the improved bias stability achieved when a source resistor R_S is used. Specifically, we change (in the MOSFET level-1 model) the value of the zero-bias threshold voltage parameter VTO by $\pm$ 0.1 V and perform bias-point simulation in Multisim. Table B.14 shows the corresponding variations in I_D and V_D for the case in which $R_S = 1.39\ \text{k}\Omega$. For the case without source degeneration, we use an $R_S = 0$ in the given schematic. Furthermore, to obtain the same I_D and V_D in both cases (for the nominal threshold voltage $V_{t0} = 0.45$ V), we use $R_{G1} = 1.2\ \text{M}\Omega$ and $R_{G2} = 0.6\ \text{M}\Omega$.

Table B.14 Variations in VTO

VTO (V)	I_D (μA)	I_D % Change	V_D (V)	V_D % Change
		With $R_S = 1.39\ \text{k}\Omega$		
0.45	250	0	0.948	0
0.35	309	23.60%	0.748	−21.10%
0.55	192	−37.86%	1.14	20.25%
		Without R_S		
0.45	255.96	0	0.9292	0
0.35	492	96.80%	0.122	−87.13%
0.55	30.1	−90.26%	1.7	127.27%

Also, Table B.15 shows the worst-case deviation of I_D and V_D values, when imposing 5% tolerance on the resistors that determine the gate voltage.

Table B.15 Variations Due to Resistor Tolerances

			With $R_S = 1.39$ kΩs			
	R_{G1} (MΩ)	R_{G2} (M Ω)	I_D (μA)	I_D % Change	V_D (V)	V_1 % Change
Nominal	0.8525	0.9475	250	0	947.67	0
I_D low V_D high	0.895	0.9	223.86	−10.44%	1.037	9.39%
I_D high V_D low	0.81	0.995	276.1	10.46%	0.858	−9.41%
			Without R_S			
	R_{G1} (M Ω)	R_{G2} (MΩ)	I_D (μA)	I_D % Change	V_D (V)	V_D % Change
Nominal	1.2	0.6	255.96	0	0.9292	0
I_D low V_D high	1.26	0.57	143.28	−44.02%	1.311	41.44%
I_D high V_D low	1.14	0.63	398.62	55.74%	0.447	−52.47%

Accordingly, we see that the source-degeneration resistor makes the bias point of the CS amplifier less sensitive to changes in the threshold voltage and the values of gate resistors. However, unless a large bypass capacitor C_S is used, this reduced sensitivity comes at the expense of a reduction in gain.

Largest Allowable Input Signal Swing Next, we wish to analyze this amplifier circuit to determine the largest allowable v_{sig} for which the transistor remains in saturation:

$$v_{DS} \geq v_{GS} - v_t$$

By enforcing this condition, with equality, at the point v_{GS} is maximum and v_{DS} is correspondingly minimum, we write:

$$v_{DS,\text{min}} \geq v_{GS,\text{max}} - v_{t0}$$

$$v_{DS} - |G_v| v_{\text{sig}} = V_{GS} + v_{\text{sig}} - v_{t0}$$

$$v_{\text{sig}} = \frac{V_{DS} - V_{GS} + V_{t0}}{(1 + |G_v|)} = \frac{0.9475 - 0.6 + 0.45}{11} = 72.5\,\text{mV}$$

This can be verified from Ch5_CS_Amplifier_Ex_swing.ms10 simulation setup. If we increase the source signal's amplitude beyond approximately 73 mV, we can observe the distortion in the output signal, indicating that the MOSFET has entered the triode region.

Amplifier Linearity Finally, we can investigate the linearity of the designed amplifier. To do so, we use the setup in Ch5_CS_Amplifier_Ex_linearity.ms10. In this case, we use a triangular waveform and increase the amplitude of the signal until the output waveform begins to show nonlinear distortion (i.e., the rising and falling edges are no longer straight lines). Based on hand analysis, linearity holds as long as $v_{\text{in}} \ll 2V_{ov}$. According to the simulation results, linearity holds until v_{in} reaches the value of approximately 30 mV, which is one-tenth of the value of $2V_{ov}$.

Example MS.7.2

Dependence of β on the Bias Current

In this example, we use Multisim to investigate the dependence of β_{dc} on the collector bias current of the Q2N3904 discrete BJT (from Fairchild Semiconductor) whose model parameters are listed in Table B.16 and are available in Multisim. As shown in the schematic capture of Fig. B.73, the V_{CE} of the BJT is fixed using a constant voltage source (in this example, $V_{CE} = 2$ V) and a dc current source I_B is applied at the base. To obtain the dependence of β_{dc} on the collector current I_C, we perform a dc-analysis simulation in which the sweep variable is the current source I_B. The β_{dc} of the BJT, which corresponds to the ratio of the collector current I_C to the base current I_B, can then be plotted versus I_C (by exporting the data to a graphing software), as shown in Fig. B.74. We see that to operate at the maximum value of β_{dc} (i.e., $\beta_{dc} = 163$), at $V_{CE} = 2$ V, the BJT must be biased at an $I_C = 10$ mA. Since increasing the bias current of a transistor increases the power dissipation, it is clear from Fig. B.74 that the choice of current I_C is a trade-off between the current gain β_{dc} and the power dissipation. Generally speaking, the optimum I_C depends on the application and technology in hand. For example, for the Q2N3904 BJT operating at $V_{CE} = 2$ V, decreasing I_C by a factor of 20 (from 10 mA to 0.5 mA) results in a drop in β_{dc} of about 25% (from 163 to 123).

Table B.16 SPICE Model Parameters of the Q2N3904 Discrete BJT

Is = 6.734f	Bf = 416.4	Xtb = 1.5	Ikr = 0	Vjc = .75	Vje = .75	Vtf = 4
Xti = 3	Ne = 1.259	Br = .7371	Rc = 1	Fc = .5	Tr = 239.5n	Xtf = 2
Eg = 1.11	Ise = 6.734f	Nc = 2	Cjc = 3.638p	Cje = 4.493p	Tf = 301.2p	Rb = 10
Vaf = 74.03	Ikf = 66.78m	Isc = 0	Mjc = .3085	Mje = .2593	Itf = .4	

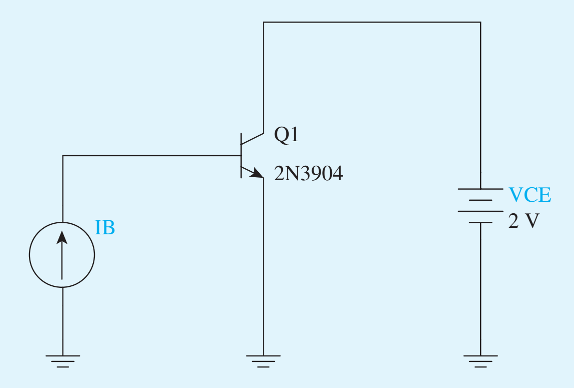


Figure B.73 The test bench used to investigate the dependence of β_{dc} on the bias current for the Q2N3904 discrete BJT.

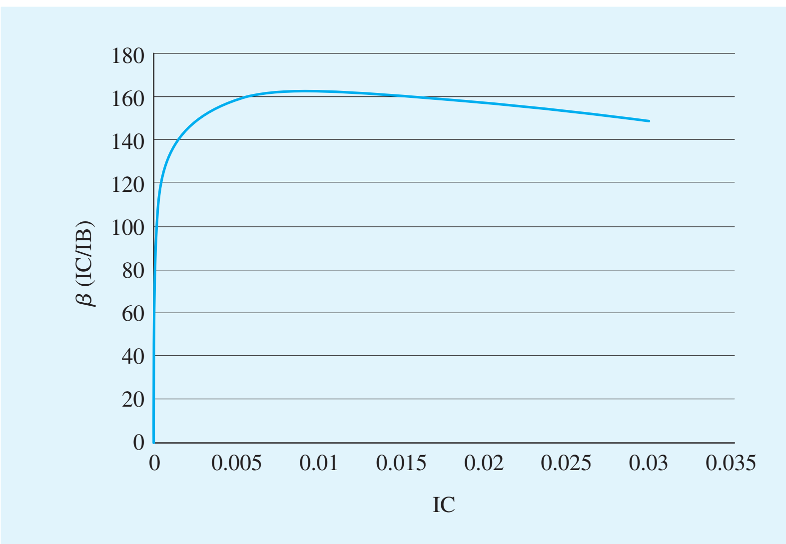


Figure B.74 Dependence of β_{dc} on I_C (at $V_{CE} = 2$ V) in the Q2N3904 discrete BJT.

Example MS.7.3

The CE Amplifier with Emitter Resistance

In this example, we use Multisim to compute the voltage gain and frequency response of the CE amplifier and investigate its bias-point stability. A schematic capture of the CE amplifier is shown in Fig. B.75. We will use part Q2N3904 for the BJT and a ±5-V power supply. We will assume a signal-source resistor $R_{\text{sig}} = 10$ kΩ, a load resistor $R_L = 10$ kΩ, and bypass and coupling capacitors of 10 μF. To enable us to investigate the effect of including a resistance in the signal path of the emitter, a resistor R_{ce} is connected in series with the emitter bypass capacitor C_E. Note that the roles of R_E and R_{ce} are different. Resistor R_E is the dc emitter-degeneration resistor because it appears in the dc path between the emitter and ground. It is therefore used to help stabilize the bias point for the amplifier. The equivalent resistance $R_e = R_E \| R_{ce}$ is the small-signal emitter-degeneration resistance because it appears in the ac (small-signal) path between the emitter and ground and helps stabilize the gain of the amplifier. In this example, we will investigate the effects of both R_E and R_e on the performance of the CE amplifier. However, as should always be the case with computer simulation, we will begin with an approximate hand analysis. In this way, maximum advantage and insight can be obtained from simulation.

Based on the plot of β_{dc} versus I_B in Fig. B.74, a collector bias current I_C (i.e., $\beta_{\text{dc}} I_B$) of 0.5 mA is selected for the BJT, resulting in $\beta_{\text{dc}} = 123$. This choice of I_C is a reasonable compromise between power dissipation and current gain. Furthermore, a collector bias voltage V_C of 0 V (i.e., at the mid–supply rail) is selected to achieve a high signal swing at the amplifier output. For $V_{CE} = 2$ V, the result is that $V_E = -2$ V,

Example MS.7.3 *continued*

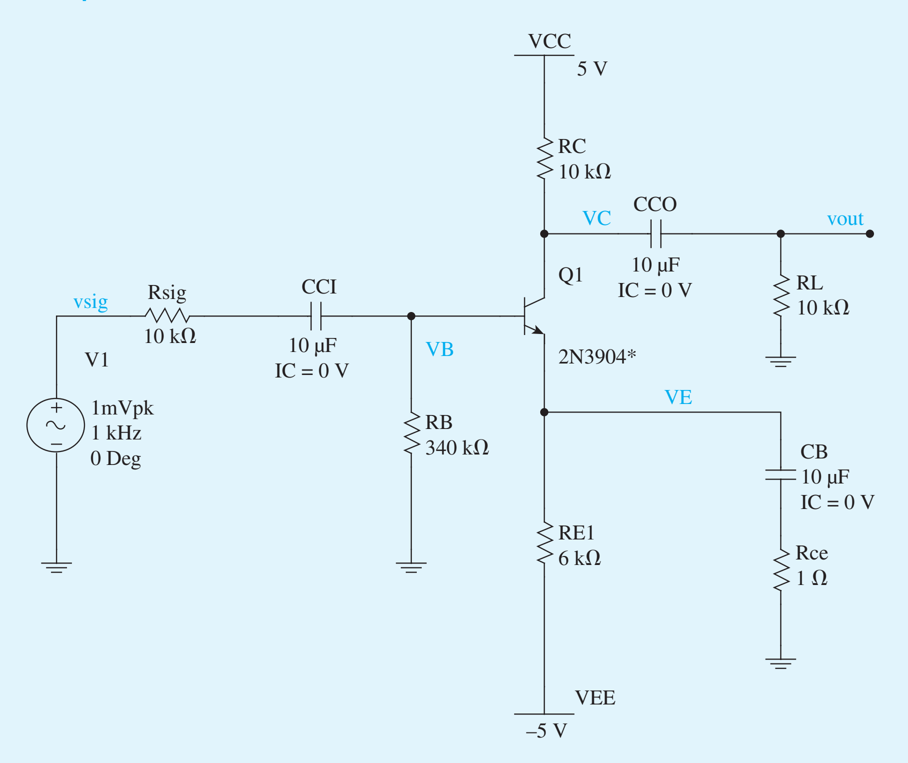


Figure B.75 Schematic capture of the CE amplifier.

requiring bias resistors with values

$$R_C = \frac{V_{CC} - V_C}{I_C} = 10\,\text{k}\Omega$$

$$R_E = \frac{V_E - V_{EE}}{I_C} = 320\,\text{k}\Omega$$

Assuming $V_{BE} = 0.7$ V and $\beta_{\text{dc}} = 123$, we can determine

$$R_B = -\frac{V_B}{I_B} = -\frac{V_{BE} + V_E}{I_C/\beta_{dc}} = 320\,\text{k}\Omega$$

Next, the input resistance R_{in} and the voltage gain $|A_v|$ of the CE amplifier:

$$R_{\text{in}} = R_B\,(\beta_{\text{ac}} + 1)(r_e + R_e)$$

$$|A_v| = \left| -\frac{R_{\text{in}}}{R_{\text{sig}} + R_{\text{in}}} \times \frac{R_C \,\|\, R_L}{r_e + R_e} \right|$$

For simplicity, we will assume $\beta_{\text{ac}} \simeq \beta_{\text{dc}} = 123$, resulting in

$$r_e = \left(\frac{\beta_{\text{ac}}}{\beta_{\text{ac}} + 1}\right)\left(\frac{V_T}{I_c}\right) = 49.6\,\Omega$$

Thus, with no small-signal emitter degeneration (i.e., $R_{ce} = 0$), $R_{\text{in}} = 6.1\ \text{k}\Omega$ and $|A_v| = 38.2$ V/V. Using the equation found for $|A_v|$ and assuming that R_B is large enough to have a negligible effect on R_{in}, it can be shown that the emitter-degeneration resistance R_e decreases the voltage gain $|A_v|$ by a factor of

$$\frac{1 + \dfrac{R_e}{r_e} + \dfrac{R_{\text{sig}}}{r_\pi}}{1 + \dfrac{R_{\text{sig}}}{r_\pi}}$$

Therefore, to limit the reduction in voltage gain to a factor of 2, we will select

$$R_e = r_e + \frac{R_{\text{sig}}}{\beta_{\text{ac}} + 1}$$

Thus, $R_{ce} \simeq R_e = 130\ \Omega$. Substituting this value in the equations found for $|A_v|$ and R_{in} shows that R_{in} increases from 6.1 kΩ to 20.9 kΩ while $|A_v|$ drops from 38.2 V/V to 18.8 V/V.

We will now use Multisim to verify our design and investigate the performance of the CE amplifier. We begin by performing a bias-point simulation to verify that the BJT is properly biased in the active region and that the dc voltages and currents meet the desired specifications. Based on this simulation forward, we have increased the value of R_B to 340 kΩ in order to limit I_C to about 0.5 mA while using a standard 1% resistor value. Next, to measure the gain A_v, we conduct a transient response analysis, as set up in Ch6_CE_Amplifier_Ex.ms10. Accordingly, with no emitter degeneration, the gain is $|A_v| = 38.5$ V/V. Using $R_{ce} = 130\ \Omega$ results in a drop in the gain by a factor of 2 (as can be seen from Fig. B.76).

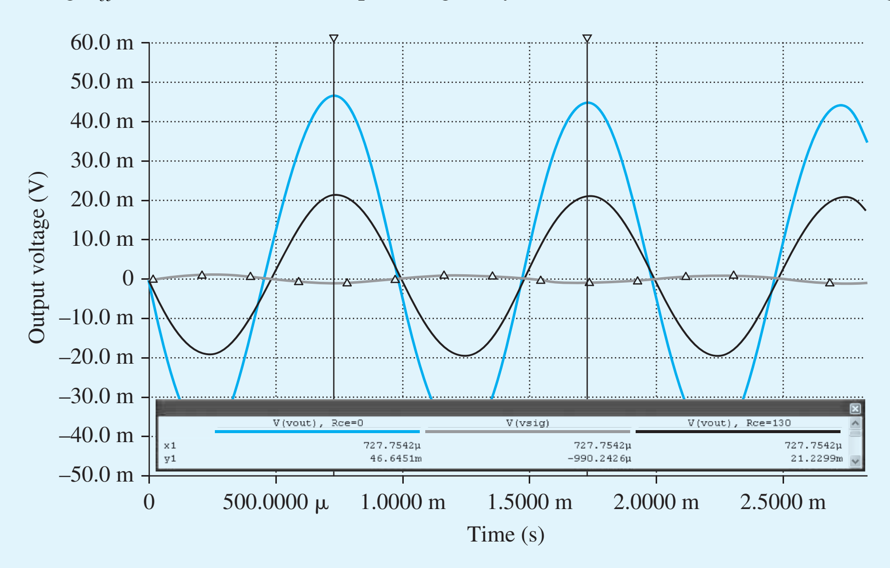


Figure B.76 Transient analysis of the CE amplifier with $R_{ce} = 0$ and $R_{ce} = 130\ \Omega$.

Thus far in this example, we have assumed that the voltage gain of the BJT amplifier is constant and independent of the frequency of the input signal. However, as mentioned in Section 10.6, this is not true, since it implies that the amplifier has infinite bandwidth. To illustrate the finite bandwidth, we compute the frequency response of the amplifier. The plot of the output-voltage magnitude (in dB) versus frequency is

Example MS.7.3 *continued*

shown in Fig. B.77. With no emitter degeneration, the midband gain is $|A_M| = 38.5$ V/V $= 31.7$ dB and the 3-dB bandwidth is $BW = f_H - f_L = 145.7$ kHz. Using an R_{ce} of 130 Ω results in a drop in the midband gain $|A_M|$ by a factor of 2 (consistent with what we observed previously in our transient analysis). Interestingly, however, BW has now increased by approximately the same factor as the drop in $|A_M|$. As we learned in Chapter 11, the emitter-degeneration resistor R_{ce} provides negative feedback, which allows us to trade off gain for other desirable properties such as a larger input resistance and a wider bandwidth.

To conclude this example, we will demonstrate the improved bias-point (or dc operating point) stability achieved when an emitter resistor R_E is used. Specifically, we will increase/decrease the value of the parameter BF (i.e., the ideal maximum forward current gain) in the SPICE model for part Q2N3904 by a factor of 2 and read the bias-point probes. The corresponding change in BJT parameter (β_{dc}) and bias-point (I_C and V_{CE}) are presented in Table B.17 for the case of $R_E = 6$ kΩ.

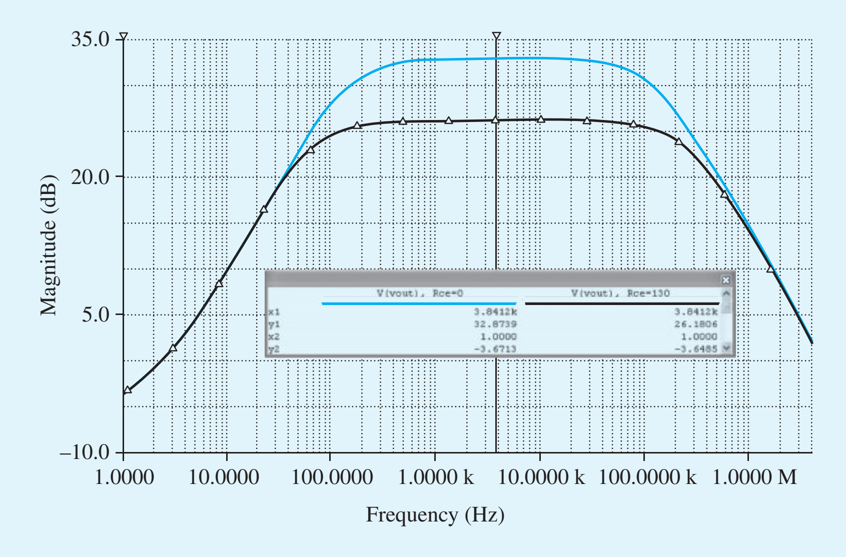


Figure B.77 Frequency response of the CE amplifier with $R_{ce} = 0$ and $R_{ce} = 130$ Ω.

Table B.17 Variations in the Bias Point of the CE Amplifier with the SPICE Model-Parameter BF of BJT

BF (in SPICE)	$R_E = 6$ kΩ			$R_E = 0$		
	β_{dc}	I_C (mA)	V_C (V)	β_{dc}	I_C (mA)	V_C (V)
208	94.9	0.452	0.484	96.9	0.377	1.227
416.4 (nominal value)	123	0.494	0.062	127	0.494	0.060
832	144	0.518	–0.183	151	0.588	–0.878

For the case without emitter degeneration, we will use $R_E = 0$ in the schematic of Fig. B.75. Furthermore, to maintain the same I_C and V_C in both cases at the values obtained for nominal BF, we use $R_B = 1.12$ MΩ to limit I_C to approximately 0.5 mA. The corresponding variations in the BJT bias

point are also shown in Table B.17. Accordingly, we see that emitter degeneration makes the bias point of the CE amplifier much less sensitive to changes in β. However, unless a large bypass capacitor C_E is used, this reduced bias sensitivity comes at the expense of a reduction in gain (as we observed in this example when we simulated the transient response of the CE amplifier with an $R_{ce} = 130\ \Omega$).

Example MS.8.1

The CMOS CS Amplifier

In this example, we will use Multisim to characterize the CMOS CS amplifier whose schematic capture is shown in Fig. B.78. We will assume a 0.18-μm CMOS technology for the MOSFET and use typical SPICE level-1 model parameters for this technology as given in Table B.4. We will begin with an approximate hand-analysis design. We will then use Multisim to investigate the performance of the final design. The targeted specifications for this CMOS CS amplifier are a voltage gain $|G_v| = 50$ V/V and a bias current I_D of 100 μA.

The amplifier specifications are summarized in Table B.18.

Table B.18 CMOS CS Amplifier Specifications

Specification	Value
I_D	100 μA
$\|G_v\|$	50 V/V
V_{DD}	1.8 V

NMOS	
Vtn	0.5 V
|VAn|	12.5 V
kn'	246.2uA/V^2
I	0.1 mA
L	0.2 um
W	0.523 um

PMOS	
Vtp	–0.5 V
|VAp|	9V
kp'	–86.1 uA/V^2
I	0.1 mA
L	0.2 um
W	0.46 um

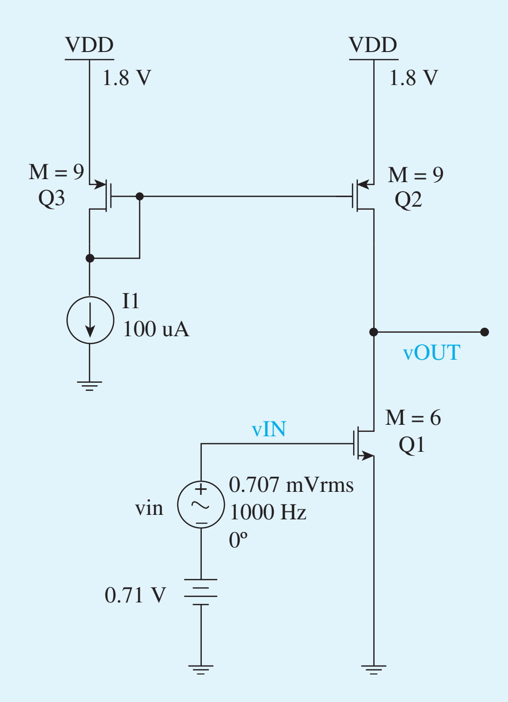


Figure B.78 Schematic capture of the CMOS CS amplifier.

Example MS.8.1 *continued*

Hand Design

For the design of this amplifier we choose $L = 0.20\,\mu\text{m}$, so that similar to Example MS.7.1, we have $L_{\text{eff}} = 0.18\,\mu\text{m}$. For this channel length, and in 0.18-μm CMOS technology, the magnitudes of the Early voltages of the NMOS and PMOS transistors are $V_{An} = 12.5$ V and $|A_{Ap}| = 9$ V, respectively. Therefore, the value of V_{OV1} can now be calculated as follows:

$$G_v = -g_m R_L' = -g_m(r_{o1} \parallel r_{o2}) = -\frac{2}{V_{OV1}}\left(\frac{V_{An}|V_{Ap}|}{V_{An}|V_{Ap}|}\right)$$

$$V_{OV1} = -\frac{2}{G_v}\left(\frac{V_{An}|V_{Ap}|}{V_{An}|V_{Ap}|}\right) = -\frac{2}{(-50)}\left(\frac{12.5 \times 9}{12.5 + 9}\right) \simeq 0.21\,\text{V}$$

MOSFET 1 can now be sized (by ignoring the channel-length modulation) as

$$\frac{W_1}{L_{\text{eff}}} = \frac{I_D}{\frac{1}{2}k_n' V_{OV1}^2} = \frac{100 \times 10^{-6}}{\frac{1}{2} \times 246.2 \times 10^{-6} \times 0.21^2} \simeq 18.42$$

where, as mentioned, $L_{\text{eff}} = 0.180\,\mu\text{m}$, and similar to Example MS.7.1, $k_n' = 246.2\,\mu\text{A/V}^2$. This yields $W_1 = 18.42\ L_{\text{eff}} = 3.32\,\mu\text{m}$. To specify the dimensions of the MOSFETs in Multisim, we will use the multiplicative factor m; its default value is 1, and it is used in SPICE to specify the number of MOSFETs connected in parallel. As depicted in Fig. B.79, a transistor with channel length L and channel width $m \times W$ can be implemented using m narrower transistors in parallel, each having a channel length L and a channel width W. In this example, a unit-size NMOS transistor is used with $W_1/L_1 = 0.52\,\mu\text{m}/0.2\,\mu\text{m}$. Thus, we find $m_1 = 3.32/0.52 \simeq 6$.

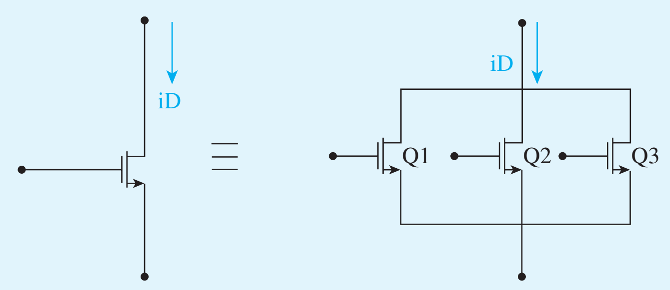


Figure B.79 Transistor equivalency.

Furthermore, MOSFETs 2 and 3 must be sized to have reasonably small V_{OV} for the bias current I_D of 100 μA. This allows large signal-swing at the output of the amplifier. Similar to our previous approach, by choosing $|V_{OV2}| = 0.3$ V, and noting $|D_{DS,2}| \simeq (V_{DD}/2) = 0.9\,\text{V}$ (mid–rail voltage):

$$\frac{W_2}{L_{\text{eff}}} = \frac{I_D}{\frac{1}{2}|k_p'|V_{OV2}^2(1 + |\lambda_2||V_{DS,2}|)} = \frac{100 \times 10^{-6}}{\frac{1}{2} \times 86.1 \times 10^{-6} \times 0.3^2 \times (1 + 0.11 \times 0.9)} \simeq 23.5$$

where $L_{eff} = 0.18\,\mu m$ and $|k'_p| = 86.1\,\mu A/V^2$. This yields $W_2 = 23.5 \times L_{eff} = 4.23\,\mu m$. In this example, unit-size PMOS transistors are used with $W_2/L_2 = W_3/L_3 = 0.46\,\mu m/0.2\,\mu m$. Thus, we find $m_2 = m_3 = 4.23/0.46 \simeq 9$.

Simulation

Amplifier Biasing Now our design can be verified using the simulation tool. The schematic is in Ch7_CMOS_CS_Amplifier_Ex_VTC.ms10. Based on the simulation results, $|G_v| = 54$ V/V and $I_D = 101\,\mu A$. Therefore, the simulation results confirm that the designed CMOS CS amplifier meets the specifications.

DC Voltage Transfer Characteristic To compute the dc transfer characteristic of the CS amplifier, we perform a dc analysis in Multisim with V_{IN} swept over the range 0 to V_{DD} and plot the corresponding output voltage V_{OUT}.

Figure B.80(a) shows the resulting transfer characteristic. The slope of the VTC curve at $V_{GS1} = 0.71$ V corresponds to the desired gain of the amplifier. To examine the high-gain region more closely, we repeat the dc sweep for V_{IN} between 0.6 V and 0.8 V. The resulting transfer characteristic is plotted in Fig. B.80(b) (middle curve). Using the cursor of the Grapher in Multisim, we find that the linear region of this dc

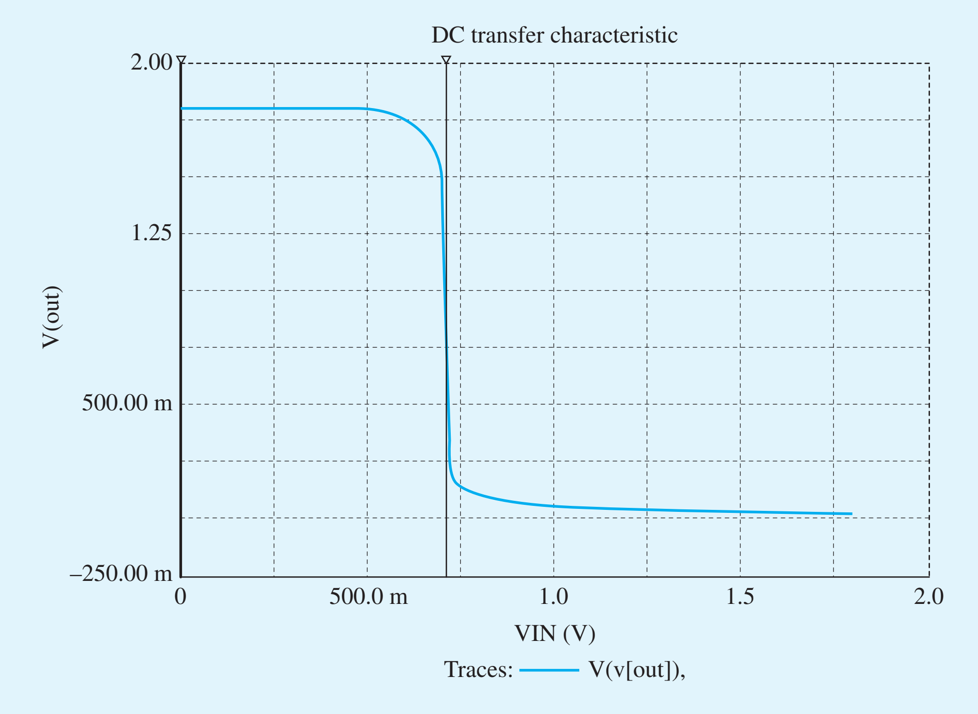


Figure B.80 **(a)** Voltage transfer characteristic of the CMOS CS amplifier.

Example MS.8.1 *continued*

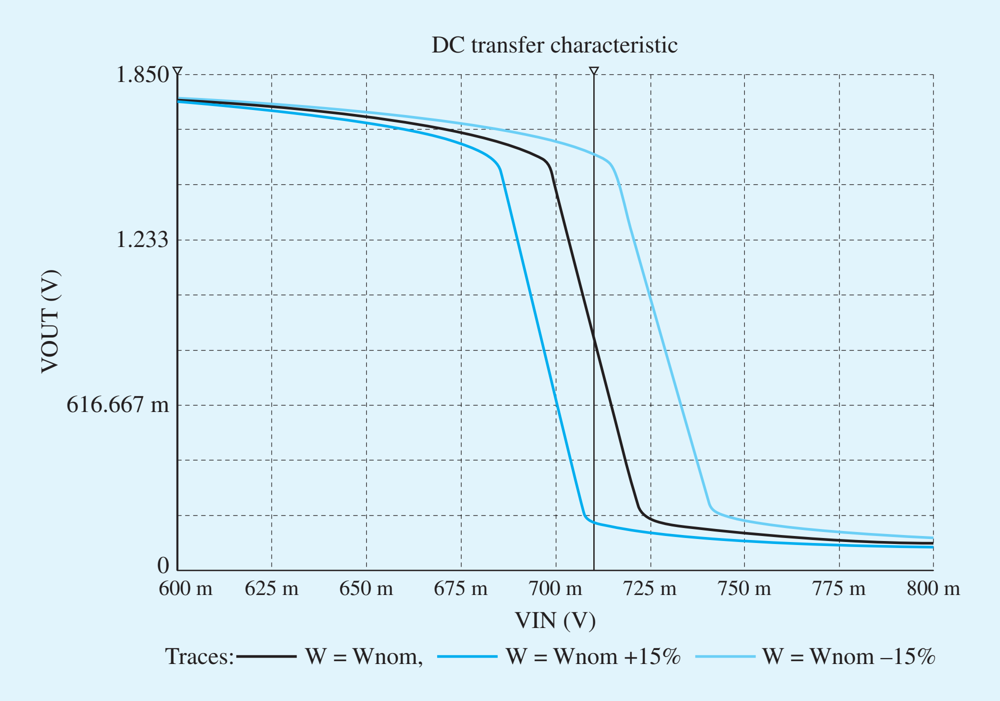


Figure B.80 **(b)** Expanded view of the transfer characteristics in the high-gain region for $W = W_{\text{nominal}} \pm 15\%$.

transfer characteristic is bounded approximately by $V_{\text{IN}} = 0.698$ V and $V_{\text{IN}} = 0.721$ V. The corresponding values of V_{OUT} are 1.513 V and 0.237 V. These results are close to the expected values. Specifically, transistors Q_1 and Q_2 will remain in the saturation region and, hence, the amplifier will operate in its linear region if $V_{OV1} \leq V_{\text{OUT}} \leq V_{DD} - |V_{OV2}|$ or $0.21\text{V} \leq V_{\text{OUT}} \leq 1.5\text{V}$. From the results above, the voltage gain G_v (i.e., the slope of the linear segment of the dc transfer characteristic) is approximately –54 V/V, which exceeds but is reasonably close to the targeted gain.

Note, from the dc transfer characteristic in Fig. 80(b), that for an input dc bias of $V_{\text{IN}} = 0.710$ V, the output dc bias is $V_{\text{OUT}} = 0.871$ V. This choice of V_{IN} maximizes the available signal swing at the output by setting V_{OUT} at approximately the middle of the linear segment of the dc transfer characteristic.

Using Transient Analysis to Verify G_v This can be done by conducting transient response analysis, as set up in Ch7_CMOS_CS_Amplifier_Ex_gain.ms10. As can be seen from Fig. B.81, $|G_v| \simeq |A_v| \simeq$ 54 V/V.

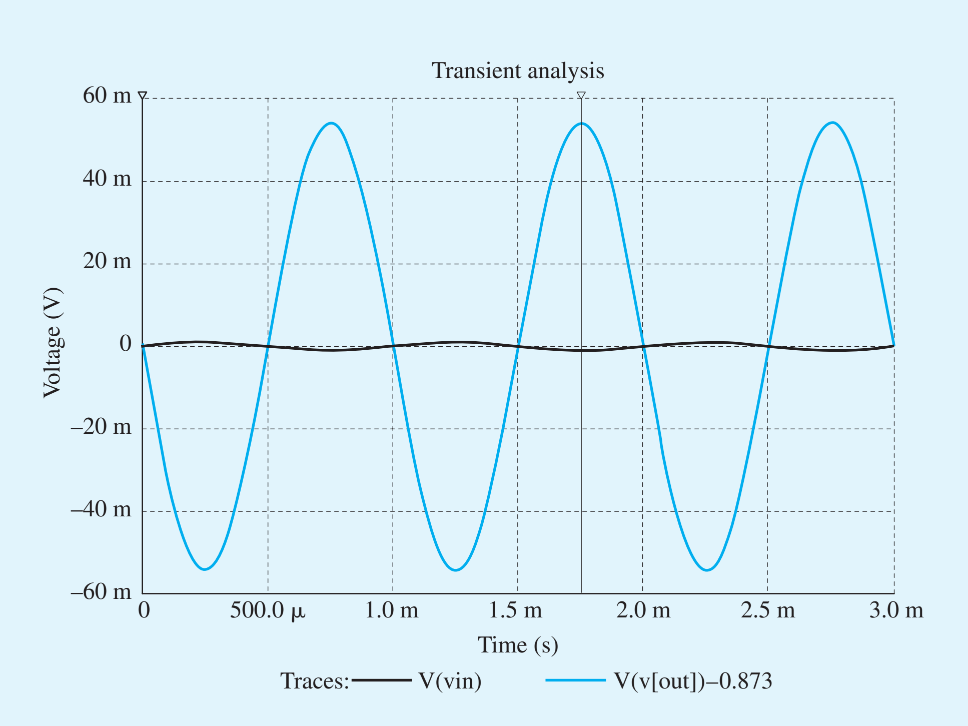


Figure B.81 G_v of the CMOS CS amplifier (transient analysis).

Sensitivity to Process Variations Because of the high resistance at the output node (or, equivalently, because of the high voltage gain), the value of V_{OUT} is highly sensitive to the effect of process and temperature variations on the characteristics of the transistors. To illustrate this point, consider what happens when the width of Q_1 (i.e., W_1) changes by ±15%. The corresponding dc transfer characteristics are shown in Fig. B.80(b). Accordingly, when $V_{IN} = 0.71$ V, V_{OUT} will drop to 0.180 V if W_1 increases by 15%, and will rise to 1.556 V if W_1 decreases by 15%. In practical circuit implementations, this problem is alleviated by using negative feedback to accurately set the dc bias voltage at the output of the amplifier and, hence, to reduce the sensitivity of the circuit to process variations. We studied the topic of negative feedback in Chapter 11.

Example MS.8.2

The Folded-Cascode Amplifier

In this example, we will use Multisim to design the folded-cascode amplifier whose schematic capture is shown in Fig. B.82. We will assume a 0.18-μm CMOS technology for the MOSFET and use typical SPICE level-1 model parameters for this technology, excluding the intrinsic capacitance values. We will begin with an approximate hand-analysis design. We will then use Multisim to verify that the designed circuit meets the specifications. The targeted specifications for this folded-cascode amplifier are a dc gain

Example MS.8.2 *continued*

$|G_v| = 100$ V/V and a bias current I_D of 100 μA. Note that while this design does not provide a very high gain, its bandwidth is large (see Chapter 10).

The amplifier specifications are summarized in Table B.19.

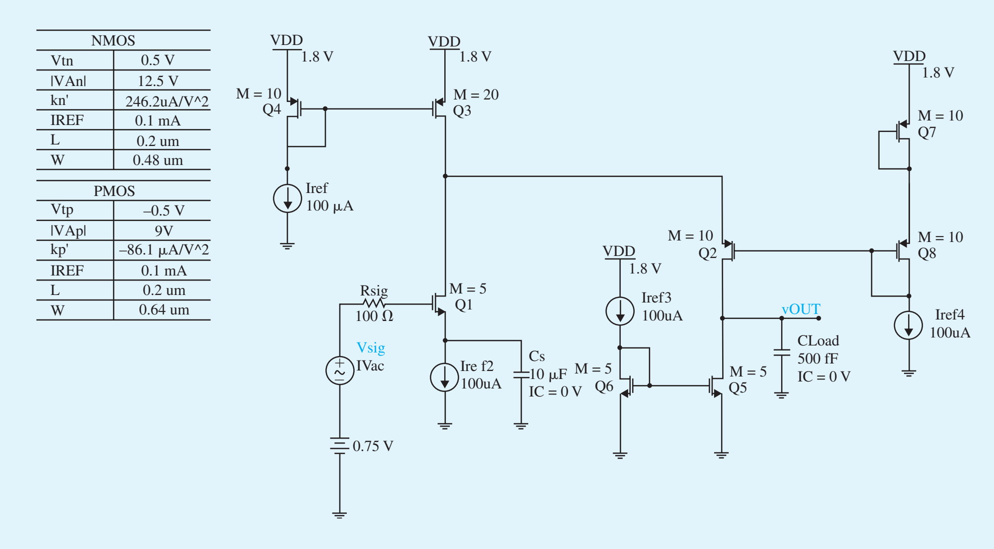


Figure B.82 Schematic capture of the folded-cascode amplifier.

Table B.19 Folded-Cascode Amplifier Specifications

Parameters	Value
I_D	100 μA
$\|G_v\|$	100 V/V
V_{DD}	1.8 V

Hand Design

For the design of this amplifier we choose $L = 200$ μm, so we have $L_{\text{eff}} = 180$ μm. For this channel length, and in 0.18-μm CMOS technology, the magnitudes of the Early voltages of the NMOS and PMOS transistors are $V_{An} = 12.5$ V and $|V_{Ap}| = 9$ V, respectively.

The folded-cascode amplifier in Fig. B.82 is equivalent to the one in Fig. 8.36, except that a current source is placed in the source of the input transistor Q_1 (for the same dc-biasing purpose as in the case of the CS amplifier). Note that in Fig. B.82, the PMOS current mirror Q_3–Q_4 and the NMOS current mirror Q_5–Q_6 are used to realize, respectively, current sources I_1 and I_2 in the circuit of Fig. 8.36. Furthermore, the current transfer ratio of mirror Q_3–Q_4 is set to 2 (i.e., $m_3/m_4 = 2$). This results in $I_{D3} \simeq 2I_{\text{ref}}$. Hence, transistor Q_2 is biased at $I_{D2} - I_{D3} - I_{D1} = I_{\text{ref}}$.

The overall dc voltage gain of the folded-cascode amplifier under design can be expressed by using Eq. (8.73) as

$$G_v = -g_{m1}R_{\text{out}}$$

where

$$R_{\text{out}} = R_{\text{out2}} \parallel R_{\text{out5}}$$

is the output resistance of the amplifier. Here, R_{out2} is the resistance seen looking into the drain of the cascode transistor Q_2, while R_{out5} is the resistance seen looking into the drain of the current mirror transistor Q_5. Using Eq. (8.93), we have

$$R_{\text{out2}} \simeq (g_{m2}r_{o2})R_{S2}$$

where

$$R_{S2} = (r_{o1} \parallel r_{o3})$$

is the effective resistance at the source of Q_2. Furthermore,

$$R_{\text{out5}} = r_{o5}$$

Thus, for the folded-cascode amplifier in Fig. B.82,

$$R_{\text{out}} \simeq r_{o5}$$

and

$$G_v = -g_{m1}r_{o5} = -2\frac{V_{An}}{V_{OV1}}$$

Therefore, based on the given information, the value of V_{OV1} can be determined:

$$V_{ov1} = 2\frac{V_{An}}{|G_v|} = 2\frac{12.5}{100} = 0.25\text{V}$$

The gate bias voltage of transistor Q_2 is generated using the diode-connected transistors Q_7 and Q_8. The size and drain currents of these transistors are set equal to those of transistor Q_2. Therefore, ignoring the body effect,

$$V_{G,2} = V_{DD} - V_{SG,7} - V_{SG,8} = V_{DD} - 2(|V_{tp}| + |V_{OV,\text{P}}|)$$

where $V_{OV,P}$ is the overdrive voltage of the PMOS transistors in the amplifier circuit. Thus, such a biasing configuration results in $V_{SG,2} = |V_{tp}| + |V_{OV,P}| = 0.5 + 0.25 = 0.75$ V as desired, while setting $V_{SD,3} = |V_{tp}| + |V_{OV,P}| = 0.75$ V to improve the bias matching between Q_3 and Q_4. For this example, all transistors are sized for an overdrive voltage of 0.25 V. Also, to simplify the design procedure, we ignore the channel-length modulation effect. As a result, using unit-size NMOS transistors with $W_n/L_n = 0.48\,\mu\text{m}/0.2\,\mu\text{m}$, and unit-size PMOS transistors with $W_p/L_p = 0.64\,\mu\text{m}/0.2\,\mu\text{m}$, the corresponding multiplicative factor m for each transistor can be calculated by rounding to the nearest integer the value of m:

$$m = \frac{I_D}{\frac{1}{2}k'\left(\frac{W}{L_{\text{eff}}}\right)V_{OV}^2}$$

Example MS.8.2 *continued*

Table B.20 summarizes the relevant design information and the calculated *m* values for each transistor.

Table B.20 Transistor Sizes

Transistor	I_D (μA)	V_{OV} (V)	W (μm)	L_{eff} (μm)	K (μA/V^2)	m
1	100	0.25	0.48	0.18	246.2	5
2	100	0.25	0.64	0.18	86.1	10
3	200	0.25	0.64	0.18	86.1	20
4	100	0.25	0.64	0.18	86.1	10
5	100	0.25	0.48	0.18	246.2	5
6	100	0.25	0.48	0.18	246.2	5
7	100	0.25	0.64	0.18	86.1	10
8	100	0.25	0.64	0.18	86.1	10

Simulation

Verifying G_v Now our design can be verified by reading probes or conducting transient response analysis, as set up in Ch7_Folded_Cascode_Ex.ms10. Based on the simulation results, $|G_v| = 102$ V/V (Fig. B.83) and $I_{D1} = I_{D2} = 100\,\mu$A. Therefore, the simulation results confirm that the designed folded-cascode amplifier meets the specifications.

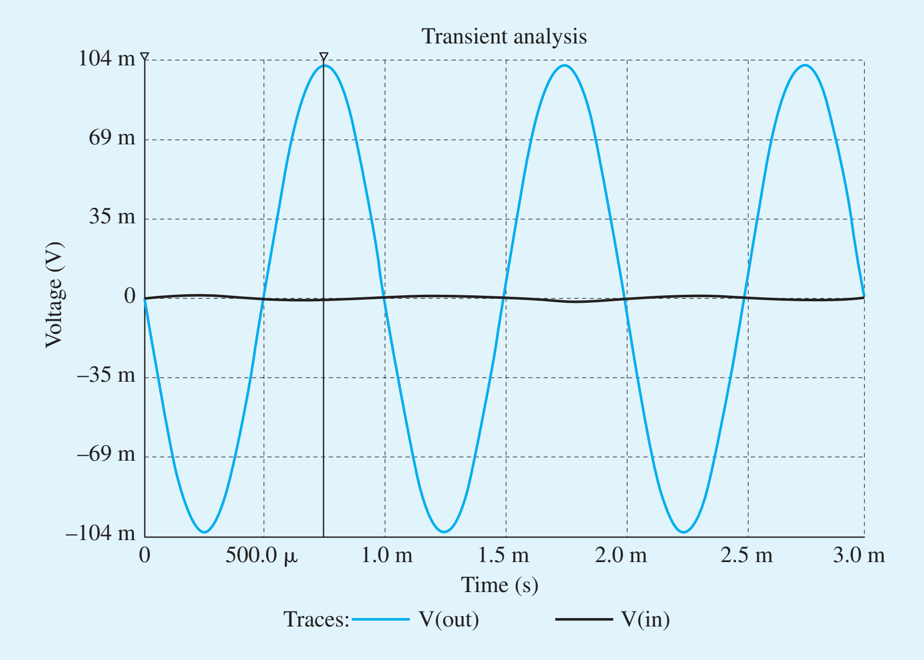


Figure B.83 G_v of the folded-cascode amplifier (transient analysis).

Sensitivity to Channel-Length Modulation In the hand design of this example, the channel-length modulation effect was ignored (except for the role of r_{o5} in determining the gain). However, the simulation took the finite r_o of each transistor into account. Furthermore, one can investigate the effect of changes in the Early voltages by modifying the value of lambda for each transistor in the design.

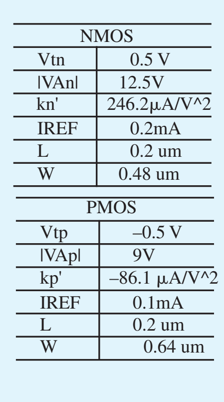
Example MS.9.1

The Two-Stage CMOS Op Amp

In this example, we will design the two-stage CMOS op amp whose schematic capture is shown in Fig. B.84. Once designed, the circuit's characteristics, such as the input common-mode range, the common-mode rejection ratio, the output-voltage range, and the input offset voltage will be evaluated.

The first stage is differential pair Q_1–Q_2 (which is actively loaded with the current mirror formed by Q_3 and Q_4), with bias current supplied by the current mirror formed by Q_8, and Q_5, which utilizes the reference bias current I_{REF}. The second stage consists of Q_6, which is a common-source amplifier actively loaded with the current source transistor Q_7.

For the design of this CMOS op amp, we will assume a 0.18-μm CMOS technology for the MOSFETs and use typical SPICE level-1 model parameters for this technology, excluding the intrinsic capacitance values. We will begin with an approximate hand-analysis design. We will then use Multisim to verify that the implemented circuit meets the specifications. The targeted specifications for this op amp are a dc open-loop voltage gain $|A_v| = 2500$ V/V, with each of transistors Q_1, Q_2, Q_3, and Q_4 conducting a drain current of 100 μA.



NMOS	
Vtn	0.5 V
|VAn|	12.5V
kn'	246.2μA/V^2
IREF	0.2mA
L	0.2 um
W	0.48 um

PMOS	
Vtp	–0.5 V
|VAp|	9V
kp'	–86.1 μA/V^2
IREF	0.1mA
L	0.2 um
W	0.64 um

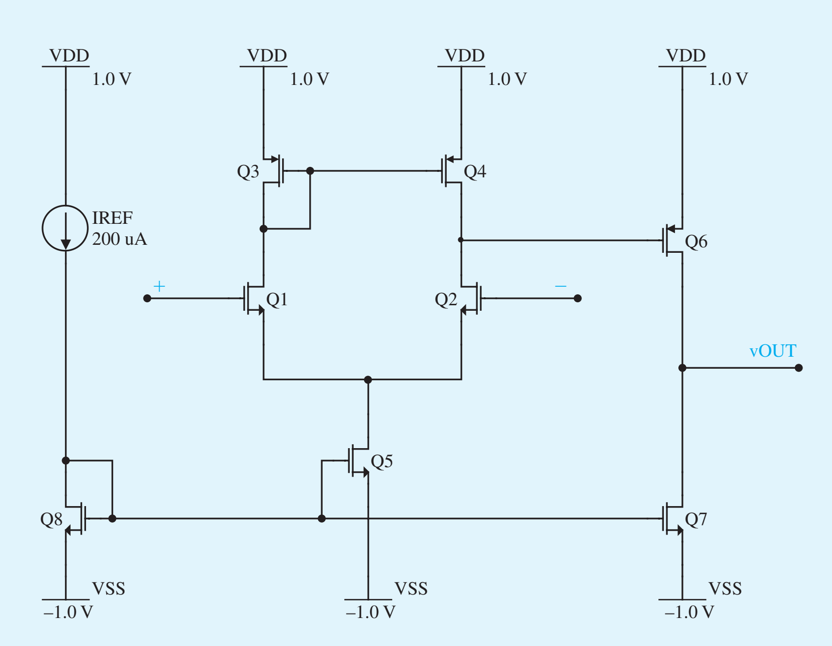


Figure B.84 Schematic capture of the two-stage CMOS op amp.

To achieve the targeted specifications, a biasing current $I_{REF} = 200$ μA is used, and the transistors Q_5, Q_6, Q_7, and Q_8 will be sized so that they conduct the drain current of 200 μA. Also, the open-loop voltage gain for this design is the product of the voltage gains of the two stages. Accordingly, each stage is designed to contribute a voltage gain of –50 V/V, so as to achieve the specified open-loop voltage gain.

The amplifier specifications are summarized in Table B.21.

Example MS.9.1 *continued*

Table B.21 Two-Stage CMOS Op-Amp Specifications

Parameter	Value
$I_{(Q1, Q2, Q3, \text{and } Q4)}$	100 μA
$I_{(Q5, Q6, Q7, \text{and } Q8)}$	200 μA
$\lvert A_1 \rvert$	50 V/V
$\lvert A_2 \rvert$	50 V/V
V_{DD}	1 V
V_{SS}	–1 V

Hand Design For the design of this amplifier we choose $L = 0.200\,\mu\text{m}$, so we have $L_{\text{eff}} = 0.180\,\mu\text{m}$. For this channel length, and in 0.18-μm CMOS technology, the magnitudes of the Early voltages of the NMOS and PMOS transistors are $V_{An} = 12.5$ V and $|V_{Ap}| = 9$ V.

The two-stage CMOS op amp in Fig. B.84 is equivalent to the one in Fig. 9.40, except that the first stage is an NMOS differential amplifier and the second stage is a PMOS common source. Note that the differential voltage gain of the first stage can be expressed using Eq. (9.37) as:

$$A_1 = -g_{m1}(r_{o2} \parallel r_{o4})$$

Hence,

$$A_1 = -\frac{2}{V_{OV1}}\left(\frac{V_{An}|V_{Ap}|}{V_{An} + |V_{Ap}|}\right)$$

resulting in

$$V_{OV1} = -\frac{2}{A_1}\left(\frac{V_{An}|V_{Ap}|}{V_{An} + |V_{Ap}|}\right) = -\frac{2}{(-50)}\left(\frac{12.5 \times 9}{12.5 + 9}\right) \simeq 0.21\text{ V}$$

Also, the voltage gain of the second stage is provided by Eq. (9.37) as

$$A_2 = -g_{m6}(r_{o6} \parallel r_{o7})$$

Therefore,

$$A_2 = -\frac{2}{V_{OV6}}\left(\frac{V_{An}|V_{Ap}|}{V_{An} + |V_{Ap}|}\right)$$

resulting in

$$V_{OV6} = -\frac{2}{A_2}\left(\frac{V_{An}|V_{Ap}|}{V_{An} + |V_{Ap}|}\right) = -\frac{2}{(-50)}\left(\frac{12.5 \times 9}{12.5 + 9}\right) \simeq 0.21\text{ V}$$

For this example, all transistors are sized for an overdrive voltage of 0.21 V. Furthermore, to simplify the design procedure, we ignore the channel-length modulation effect. As a result, using unit-size NMOS transistors with $W_n/L_n = 0.64\,\mu\text{m}/0.2\,\mu\text{m}$, and unit-size PMOS transistors with $W_n/L_n = 0.48\,\mu\text{m}/0.2\,\mu\text{m}$, the corresponding multiplicative factor m for each transistor can be calculated by rounding to the nearest

integer value which is calculated as m:

$$m = \frac{I_D}{\frac{1}{2}k'\left(\frac{W}{L_{\text{eff}}}\right)V_{OV}^2}$$

Table B.22 summarizes the relevant information and the calculated m values for each transistor.

Table B.22 Transistor Sizes

Transistor	I_D (μA)	V_{OV} (V)	W (μm)	L_{eff} (μm)	k' (μA/V^2)	m
1	100	0.21	0.48	0.18	246.2	7
2	100	0.21	0.48	0.18	246.2	7
3	100	0.21	0.64	0.18	86.1	15
4	100	0.21	0.64	0.18	86.1	15
5	200	0.21	0.48	0.18	246.2	14
6	200	0.21	0.64	0.18	86.1	30
7	200	0.21	0.48	0.18	246.2	14
8	200	0.21	0.48	0.18	246.2	14

Simulation

Verifying A_v Now our design can be verified by reading probes, as set up in Ch8_Two_Stage_Op_Amp_Ex.ms10. Based on the simulation results we read $|A_1| = 57$ V/V, $|A_2| = 58.6$ V/V, $|A_v| = 3340$ V/V, $I_{(Q1,Q2,Q3,\text{and}Q4)} = 97$ μA, $I_{Q5} = 194$ μA, $I_{(Q6,Q7)} = 202$ μA, and $I_{Q8} = 200$ μA. These values are somewhat different from the targeted specifications. The deviations can be attributed to the fact that we rounded the values of m to the nearest integer and ignored the effect of channel-length modulation, that is, the term $(1 + \lambda V_{DS})$, when calculating the multiplicative factor. To get closer to our targeted specifications, we may use the obtained V_{DS} values for each transistor, from the original design, to estimate new multiplicative factor values by taking the term $(1 + \lambda V_{DS})$ into account. Table B.23 shows the revised multiplicative factor values.

Table B.23 Revised Transistor Multiplicative Factors

Transistor	m
1	6
2	6
3	14
4	14
5	13
6	26
7	13
8	13

The revised design is evaluated by reading probes, as set up in Ch8_Two_Stage_Op_Amp_revised_Ex.ms10. The simulation results show $|A_1| = 54$ V/V, $|A_2| = 58.2$ V/V, $|A_v| = 3145$ V/V,

Example MS.9.1 *continued*

$I_{(Q1,Q2,Q3 \text{ and } Q4)} = 103\,\mu\text{A}$, $I_{Q5} = 206\,\mu\text{A}$, $I_{(Q6,Q7)} = 205\,\mu\text{A}$, and $I_{Q8} = 200\,\mu\text{A}$, from which we see that the voltage gains are closer to the targeted specifications.

One should note that the discrepancies between the hand-design and simulation results in this simulation example are more apparent because errors in each stage add up.

Next, we will explore some important characteristics of the designed two-stage CMOS op amp.

Input Common-Mode Range The upper limit of the input common-mode range is the value of input voltage at which Q_1 and Q_2 leave the saturation region. This occurs when the input voltage exceeds the drain voltage of Q_1 by $V_{tn} = 0.5$ V. Since the drain of Q_1 is at $1 - (0.21 + 0.5) = 0.29$ V, then the upper limit of the input common-mode range is $v_{ICM\max} = 0.29 + 0.5 = 0.79\,\text{V}$.

The lower limit of the input common-mode range is the value of input voltage at which Q_5 leaves the saturation region. Since for Q_5 to operate in saturation the voltage across it (i.e., V_{DS5}) should at least be equal to the overdrive voltage at which it is operating (i.e., 0.21 V), the highest voltage permitted at the drain of Q_5 should be –0.79 V. It follows that the lowest value of v_{ICM} should be $v_{ICM\min} = -0.08\,\text{V}$.

To verify the results using the simulation tool, we swept the input common-mode voltage v_{ICM} from –1 V to 1 V and plotted the resulting v_{GD} of Q_1 and Q_5 (as set up in Ch8_Two_Stage_Op_Amp_Ex_CM_Range.ms10). As can be seen from Fig. B.85, both transistors Q_1 and Q_5 stay in saturation for the input common-mode range of $-0.08\text{V} \le v_{ICM} \le 0.79\,\text{V}$, as indicated by cursors.

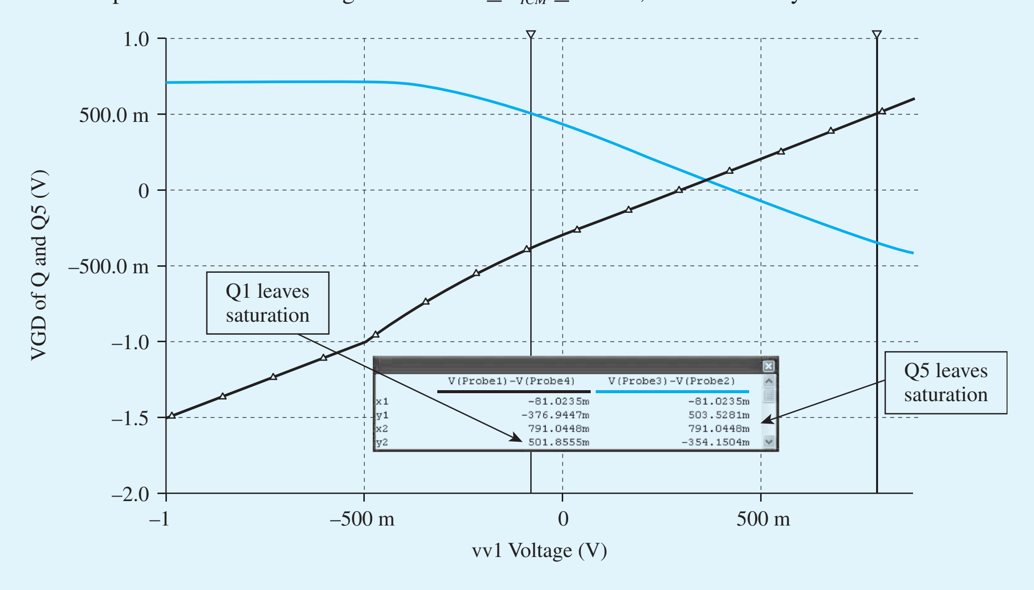


Figure B.85 Input common–mode range of the two-stage CMOS op amp.

Common-Mode Rejection Ratio (CMRR) of the First Stage The value of the CMRR of the first stage (the active-loaded MOS differential amplifier) is determined from Eq. (8.147). Note that the value of R_{SS} in the provided equation corresponds to the output resistance of Q_5 (i.e., r_{o5}). Thus,

$$\text{CMRR} \equiv \frac{|A_1|}{|A_{cm}|} = \frac{50}{1/2g_{m3}r_{o5}} = 100g_{m3}r_{o5} = 100\frac{2 \times 100 \times 10^{-6}}{0.21}\,\frac{12.5}{200 \times 10^{-6}} = 5952.4 = 75.5\text{ dB}$$

Using the simulation tool, the value of CMRR is calculated by dividing the previously obtained A_1 value (54 V/V) by the common-mode gain of the first stage as measured in Ch8_Two_Stage_Op_Amp_Ex_CMRR.ms10. This yields

$$\text{CMRR} \equiv \frac{|A_1|}{|A_{cm}|} = \frac{54}{78 \times 10^{-3}} = 6923 = 76.8\,\text{dB}$$

Output Voltage Range The lowest allowable output voltage is the value at which Q_7 leaves the saturation region, which is $-V_{SS} + V_{OV7} = -1 + 0.21 = 0.79\,\text{V}$. The highest allowable output voltage is the value at which Q_6 leaves saturation, which is $V_{DD} - |-V_{OV6}| = 1 - 0.21 = 0.79\,\text{V}$. Thus, the output-voltage range is –0.79 V to 0.79 V.

As set up in Ch8_Two_Stage_Op_Amp_Ex_Output_Range.ms10, to verify the calculated output voltage range, we swept the input voltage from –2 mV to 2 mV (we used a small input voltage due to high gain). As can be seen from Fig. B.86, the output level changes from –0.795 V to 0.784 V, a rather symmetrical range. Therefore, the simulation results confirm our hand-analysis calculations.

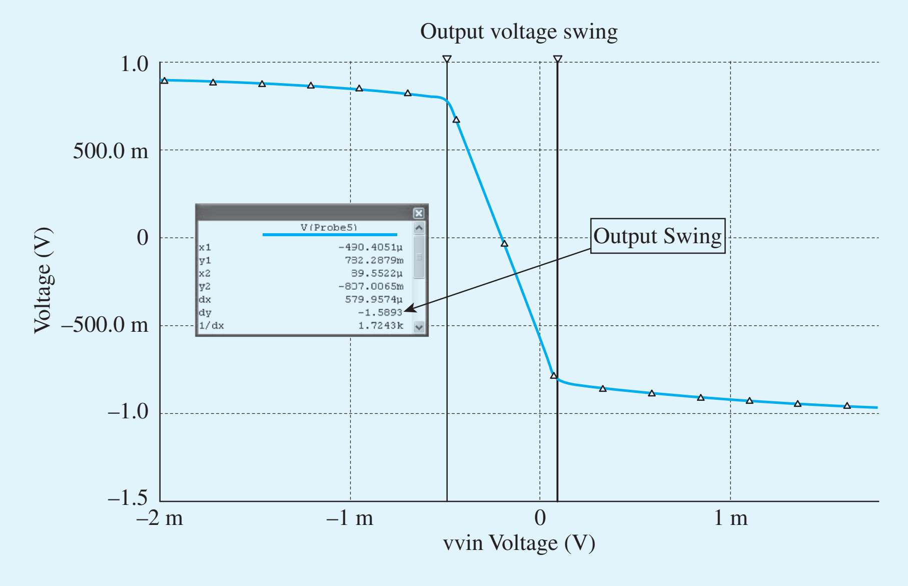


Figure B.86 Output-voltage range of the two-stage CMOS op amp.

Input Offset Voltage Although, theoretically, there should be no systematic offset, we do observe an output offset voltage V_O. As defined by Eq. 9.97, the input offset voltage, V_{OS}, can be obtained as

$$V_{OS} = \frac{V_O}{A_v}$$


Example MS.9.1 *continued*

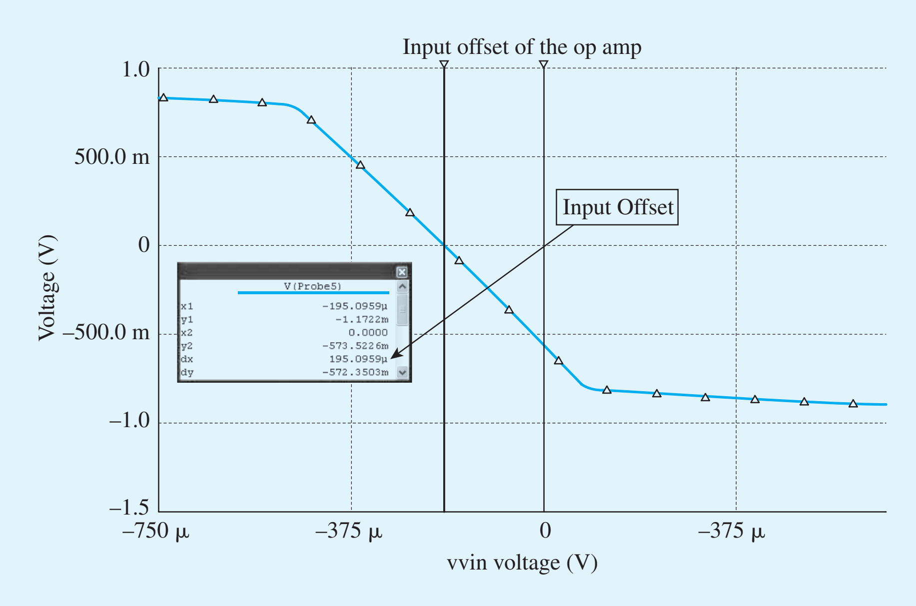


Figure B.87 Input offset voltage of the two-stage CMOS op amp.

Equivalently, if we apply a voltage $-V_{OS}$ between the input terminals of the differential amplifier, the output voltage should be reduced to zero. This equivalency can be verified using the simulation tool (Ch8_Two_Stage_Op_Amp_Ex_Output_Range.ms10). When both the input terminals are grounded, the probe at the output reads the dc voltage 0.574 V. Also, when we apply the voltage $V_{OS} = (0.574/3145) \simeq 183\,\mu\text{V}$, between the input terminals, the output voltage is reduced to zero (Fig. B.87). Hence, the op amp has an input offset voltage of $V_{OS} = 195\,\mu\text{V}$, which approximately corresponds to an output offset voltage of $V_O = 0.574$ V.

Example MS.10.1

Frequency Response of the Discrete CS Amplifier

In this example, we will investigate the frequency response of the CS amplifier of Example MS.7.1. By using Multisim to perform "ac analysis" on the designed CS amplifier, we are able to measure the midband gain A_M and the 3-dB frequencies f_L and f_H, and to plot the output-voltage magnitude (in dB) versus frequency. Figure B.88 shows the schematic capture of the CS amplifier.

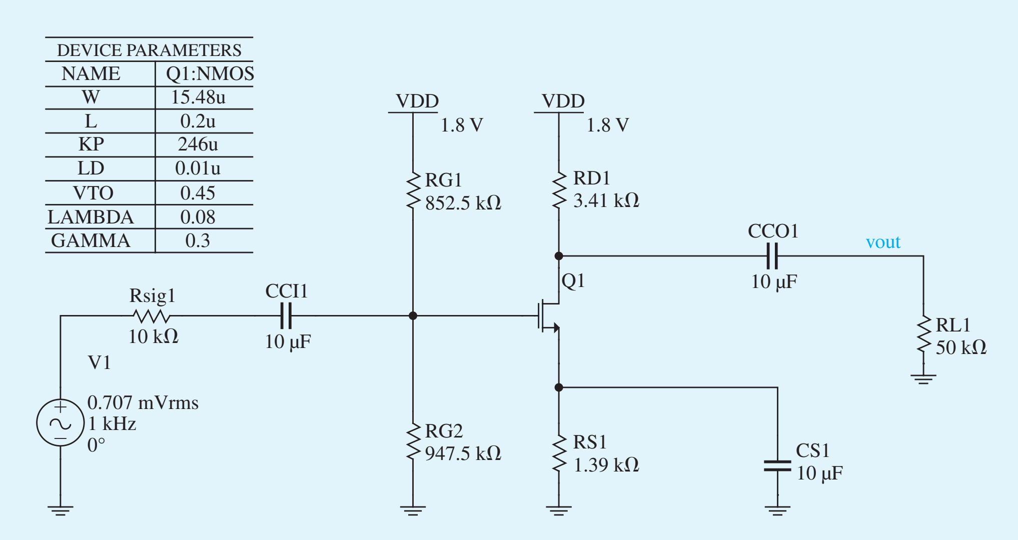


Figure B.88 Schematic capture of discrete CS amplifier.

Hand Analysis

Midband Gain The midband gain of this CS amplifier can be determined using Eq. (10.10) as follows:

$$A_M = \frac{R_{\text{in}}}{R_{\text{in}} + R_{\text{sig}}}\left[g_m\left(R_D \| R_L\right)\right]$$

$$R_{\text{in}} = \left(R_{G1} \| R_{G2}\right) = 852.5 \times 10^3 \| 947.5 \times 10^3 = 448.75 \times 10^3\ \Omega$$

$$g_m = 3.33\,\text{mA/V}$$

$$A_M = \frac{448.75 \times 10^3}{448.75 \times 10^3 + 10 \times 10^3}\left[3.33 \times 10^{-3}\left(3.41 \times 10^3 \| 50 \times 10^3\right)\right] \simeq 10\,\text{V/V}$$

Low-Frequency Poles and Zero We know from Section 10.1.1 that the low-frequency poles are as follows:

$$f_{P1} = \frac{1}{2\pi \times C_{CI}\left(R_{\text{sig}} + R_{\text{in}}\right)} = \frac{1}{2\pi \times 10 \times 10^{-6}\left[\left(10 \times 10^3\right) + 448.75 \times 10^3\right]}$$

$$f_{P1} = 0.0347\,\text{Hz}$$

$$f_{P2} = \frac{1}{2\pi \times C_{CO}\left(R_D + R_L\right)} = \frac{1}{2\pi \times 10 \times 10^{-6}\left(3.41 \times 10^3\right) + \left(50 \times 10^3\right)}$$

$$f_{P2} = 0.30\,\text{Hz}$$

$$f_{P3} = \frac{1}{2\pi \times C_S}\left(g_m + \frac{1}{R_S}\right) = \frac{1}{2\pi \times 10 \times 10^{-6}}\left[\left(3.33 \times 10^{-3}\right) + \frac{1}{1.39 \times 10^3}\right]$$

$$f_{P3} = 64.4\,\text{Hz}$$

And the location of the real transmission zero is determined as

$$f_Z = \frac{1}{2\pi \times C_S R_S} = \frac{1}{2\pi \times \left(10 \times 10^{-6}\right)\left(1.39 \times 10^3\right)}$$

$$f_Z = 11.45\,\text{Hz}$$

Example MS.10.1 *continued*

Upon observing the relative magnitude of each of the poles, we can conclude that f_{P3} will determine f_L, the lower 3-dB frequency of the amplifier gain,

$$f_L \simeq f_{P3} \simeq 11.45 \text{ Hz}$$

High-Frequency Rolloff The high-frequency rolloff of the amplifier gain is caused by the MOSFET internal capacitance. The typical values for 0.180 μm CMOS technology are given in Table B.4. We know from Eq. (10.53) in Section 10.2 that

$$f_H = \frac{1}{2\pi \times C_{\text{in}} R'_{\text{sig}}}$$

$$R'_{\text{sig}} = 10 \times 10^3 \,\|\, 448.75 \times 10^3 = 9.78 \times 10^3$$

$$C_{\text{in}} = W\{C_{gs0} + C_{gd0}[1 + g_m(R_L \,\|\, R_L)]\}$$

Note that C_{gs0} and C_{gd0} are per-unit-width values provided in the models.

$$C_{\text{in}} = (15.48 \times 10^{-6}) \times (0.3665 \times 10^{-9}) \times [1 + 1 + 3.33 \times 10^{-3}(50 \times 10^3 \,\|\, 3.41 \times 10^{-3})]$$

$$C_{\text{in}} = 0.716 \text{ fF}$$

$$f_H = \frac{1}{2\pi \times 0.716 \times 10^{-15} \times 9.78 \times 10^3}$$

$$f_H \simeq 191 \text{ MHz}$$

Now we can determine the bandwidth, BW, of the CS amplifier:

$$BW = f_H - f_L$$

$$BW \simeq f_H = 191 \text{ MHz}$$

Simulation

Figure B.89 shows the magnitude plot of the frequency response of this CS amplifier.

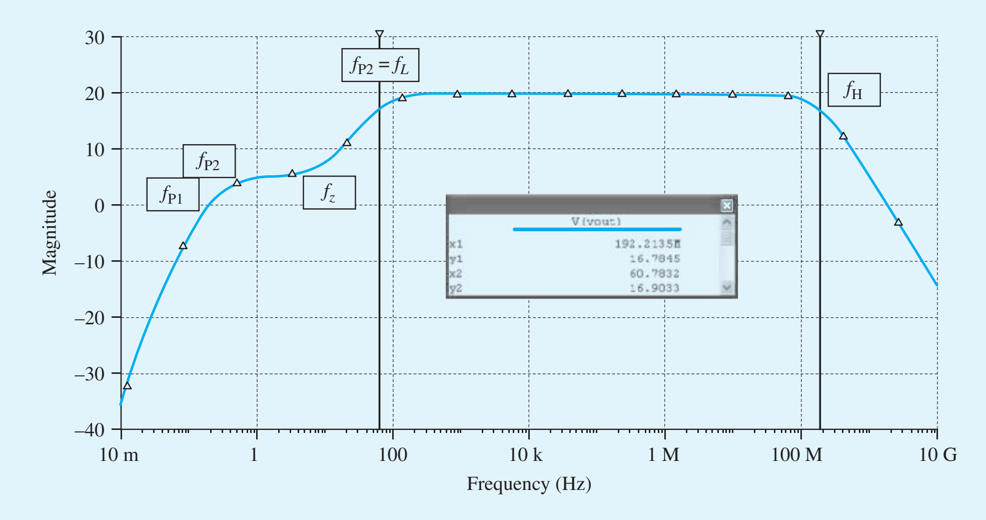


Figure B.89 Frequency response of the CS amplifier.

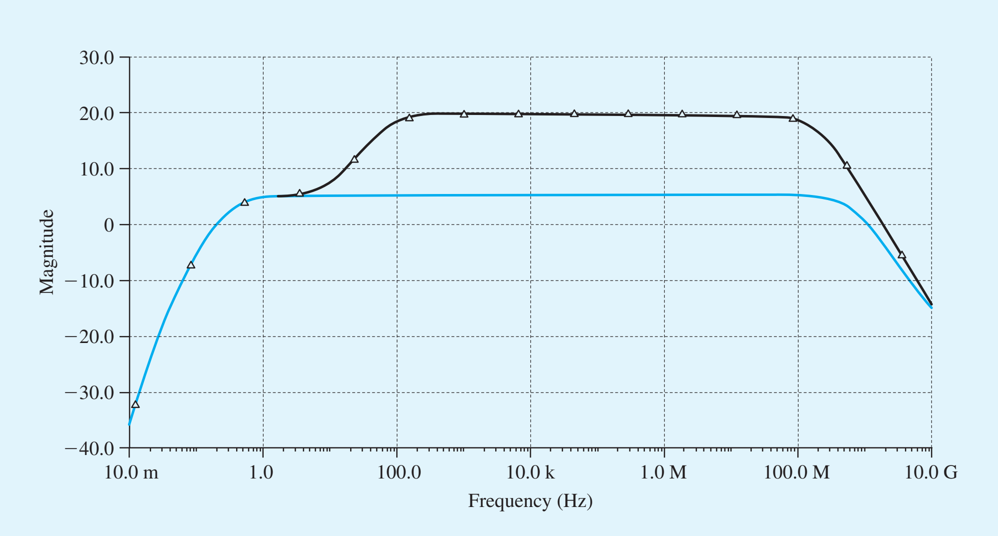


Figure B.90 Frequency response of the CS amplifier with $C_S = 10$ μF and $C_S = 0$.

Based on the simulation results, the midband gain is $A_M = 9.80$ V/V. Also, $f_L = 60.8$ Hz and $f_H =$ 192.2 MHz, resulting in 3-dB bandwidth of $BW = f_L - f_H = 192.2$ MHz. Figure B.89 further shows that (moving toward the left) the gain begins to fall off at about 300 Hz, but flattens out again at about 12.2 Hz. This flattening in the gain at low frequencies is due to a real transmission zero introduced in the transfer function of the amplifier by R_S together with C_S, with a frequency $f_Z - = 1/2\pi R_S C_S = 11.45$ Hz. Students are encouraged to investigate this relationship by using the simulation tool to modify the values of R_S and C_S and observing the corresponding change in the zero frequency. Note this value of zero is typically between the break frequencies f_{P2} and f_{P3}. The simulation is set up in Ch9_CS_Amplifier_Ex.ms10.

We can further verify this phenomenon by resimulating the CS amplifier with a $C_S = 0$ (i.e., removing C_S) in order to move f_Z to infinity and remove its effect. The corresponding frequency response is plotted in Fig. B.90. As expected with $C_S = 0$, we do not observe any flattening in the low-frequency response of the amplifier. However, because the CS amplifier now includes a source resistor R_S, the value of A_M has dropped by a factor of 5.4. This factor is approximately equal to $(1 + g_m R_S)$, as expected from our study of the CS amplifier with a source-degeneration resistance. Note that the bandwidth BW has increased by approximately the same factor as the drop in gain A_M. As we learned in Chapter 11 in our study of negative feedback, the source-degeneration resistor R_S provides negative feedback, which allows us to trade off gain for a wider bandwidth.

Example MS.10.2

The Frequency Response of CMOS CS Amplifier and the Folded-Cascode Amplifier

In this example, we will investigate the frequency response of the CMOS CS amplifier and the folded-cascode amplifier studied in Examples MS.8.1 and MS.8.2. The circuit diagram of the CMOS CS amplifier is given in Fig. B.91.

Example MS.10.2 *continued*

By using Multisim to perform "ac analysis" on the designed CMOS CS amplifier, we are able to measure the midband gain A_M and the 3-dB frequency f_H, and to plot the output-voltage magnitude (in dB) versus frequency for two different cases of R_{sig} (100 Ω and 1 MΩ), as shown in Fig. B.92.

Observe that f_H decreases when R_{sig} is increased. This is anticipated from our study of the high-frequency response of the CS amplifier. Specifically, as R_{sig} increases, the pole

$$f_{p,\text{in}} = \frac{1}{2\pi}\frac{1}{R_{\text{sig}}C_{\text{in}}}$$

formed at the amplifier input will have an increasingly significant effect on the overall frequency response of the amplifier. As a result, the effective time constant τ_H increases and f_H decreases. When R_{sig} becomes very large, as it is when $R_{sig} = 1$ MΩ, a dominant pole is formed by R_{sig} and C_{in}. This results in

$$f_H \simeq f_{p,\text{in}}$$

To estimate $f_{p,\text{in}}$, we need to calculate the input capacitance C_{in} of the amplifier. Using Miller's theorem, we have

$$C_{\text{in}} = C_{gs1} + C_{gd1}(1 + g_{m1}R'_L)$$

where

$$R'_L = r_{o1} \| r_{o2}$$

The value of C_{in} can be calculated by using the overlap capacitances $C_{gs,ov1}$ and gate-to-channel C_{gs} and $C_{gd,ov1}$ as follows:

$$C_{gs,ov_1} = m_1 W_1 CGSO = (5 \times 0.48 \times 10^{-6}) \times (0.3665 \times 10^{-9}) = 0.880\,\text{fF}$$
$$C_{gd,ov_1} = m_1 W_1 CGDO = (5 \times 0.48 \times 10^{-6}) \times (0.3665 \times 10^{-9}) = 0.880\,\text{fF}$$

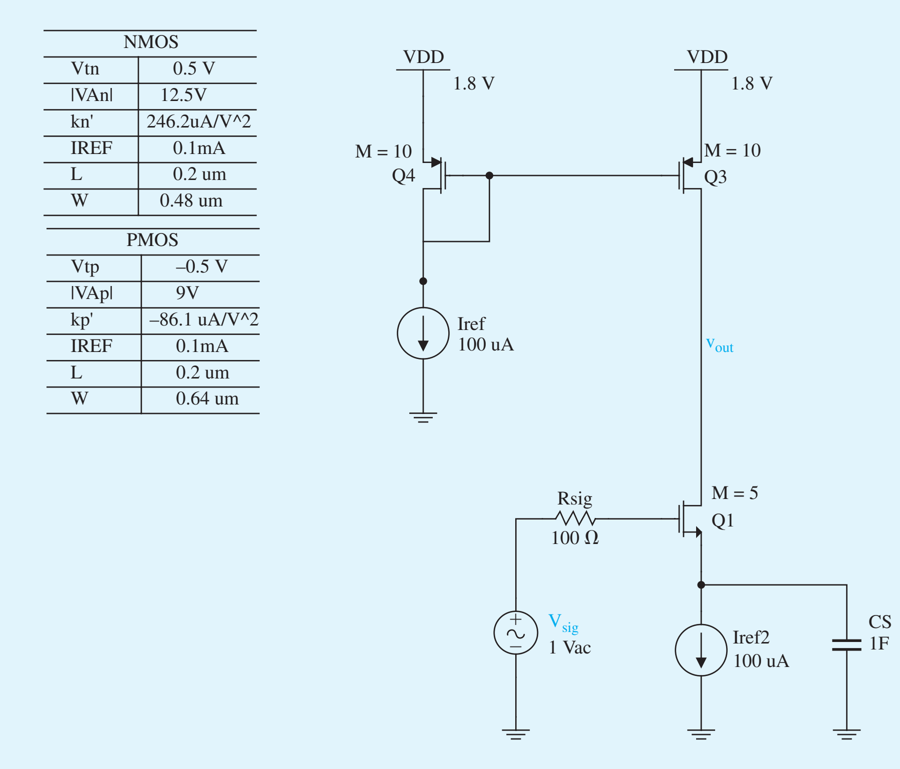
For C_{gs}, we write

$$C_{gs_\text{channel}} = \frac{2}{3}m_1 W_1 L C_{ox} = \frac{2}{3}\left[(5 \times 0.48 \times 10^{-6}) \times (0.18 \times 10^{-6})\left(\frac{3.9 \times 8.85 \times 10^{-12}}{4.08 \times 10^{-9}}\right)\right]$$
$$C_{gs_\text{channel}} = 2.48\,\text{fF}$$
$$C_{gs} = 2.48\,\text{fF} + 0.880\,\text{fF} = 3.36\,\text{fF}$$

This results in $C_{in} = 45.78$ fF when $\|G_v\| = 50$ V/V. Accordingly,

$$f_H \simeq \frac{1}{2\pi}\frac{1}{1 \times 10^6 \times 43.3 \times 10^{-15}} = 3.48\,\text{MHz}$$

which is close to the value computed by Multisim (i.e., $f_H = 3.66$ MHz).



NMOS	
Vtn	0.5 V
|VAn|	12.5V
kn'	246.2uA/V^2
IREF	0.1mA
L	0.2 um
W	0.48 um

PMOS	
Vtp	–0.5 V
|VAp|	9V
kp'	–86.1 uA/V^2
IREF	0.1mA
L	0.2 um
W	0.64 um

Figure B.91 Schematic capture of the CMOS CS amplifier.

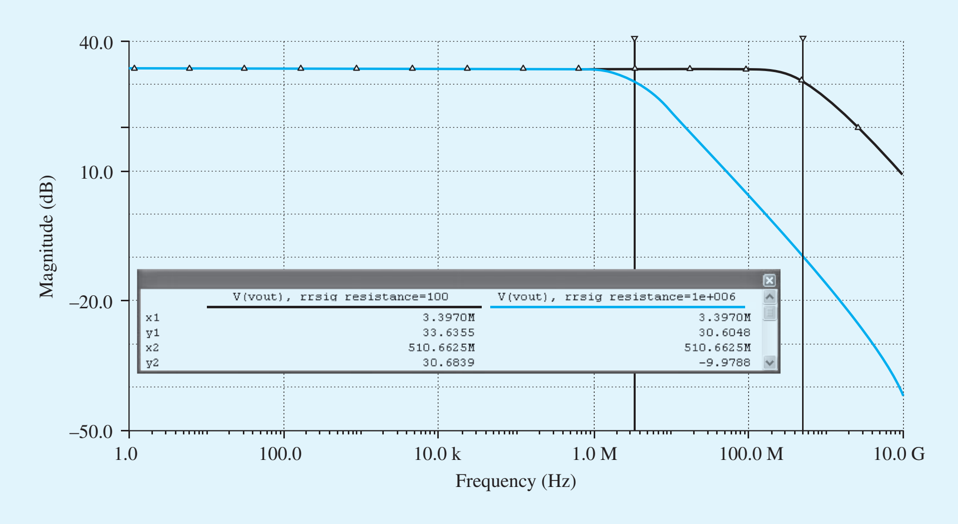


Figure B.92 Frequency response of the CMOS CS amplifier with $R_{sig} = 100\ \Omega$ and $R_{sig} = 1\ \text{M}\Omega$.

Example MS.10.2 *continued*

The Folded-Cascode Amplifier Next, we will investigate the frequency response of the folded-cascode amplifier and compare its performance with that of the CS amplifier. Figure B.93 shows the circuit diagram of the folded-cascode amplifier.

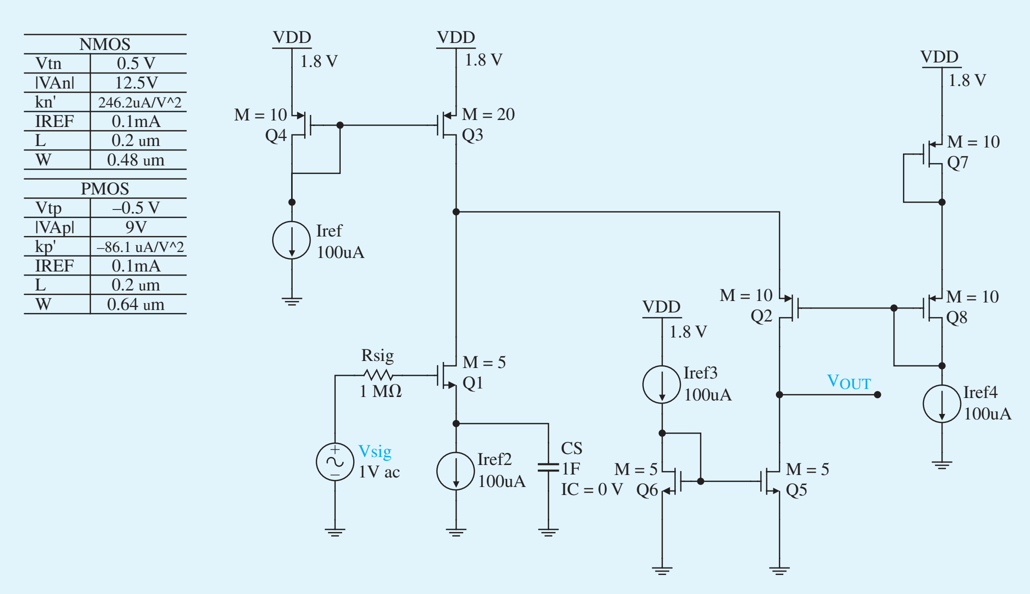


Figure B.93 Schematic capture of the folded-cascode amplifier.

Figure B.94 shows the frequency response of the folded-cascode amplifier as simulated by Multisim for the cases of $R_{sig} = 100\ \Omega$ and $1\ M\Omega$. The corresponding values of the 3-dB frequency f_H of the amplifier are given in Table B.24.

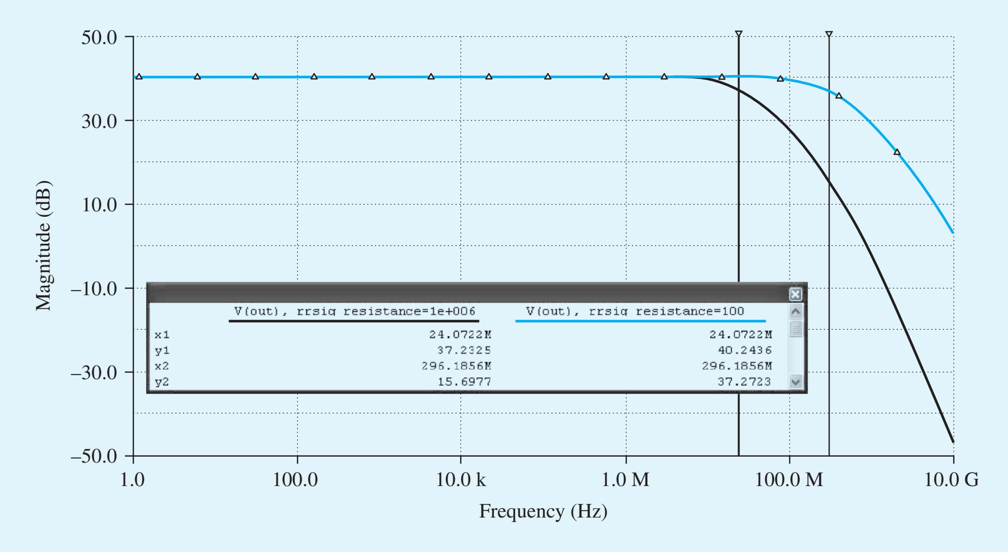


Figure B.94 Frequency response of the folded-cascode amplifier with $R_{sig} = 100\ \Omega$ and $R_{sig} = 1\ M\Omega$.

Table B.24 Dependence of f_H for the Designed Amplifiers

	f_H	
R_{sig}	CS Amplifier	Folded-Cascode Amplifier
100 Ω	510.7 MHz	296.2 MHz
1 MΩ	3.39 MHz	24.0 MHz

First, note that for the designed folded-cascode amplifier, $R_{out} = 125\ \text{k}\Omega$ and $|G_v| = 100$ V/V. Thus, R_{out} and G_v are larger than those of the CS amplifier (by a factor of 2). Note that these calculations can be found in Examples MS.8.1 and MS.8.2.

Also, observe that when R_{sig} is small, f_H of the folded-cascode amplifier is lower than that of the CS amplifier by a factor of about 1.8, approximately equal to the factor by which the gain is increased. This is because when R_{sig} is small, the frequency response of both amplifiers is dominated by the pole formed at the output node, that is,

$$f_H \simeq f_{p,\text{out}} = \frac{1}{2\pi} \frac{1}{R_{out} C_{out}}$$

Now the output resistance of the folded-cascode amplifier is larger than that of the CS amplifier, while their output capacitances are approximately equal. Therefore, the folded-cascode amplifier has a lower f_H in this case.

On the other hand, when R_{sig} is large, f_H of the folded-cascode amplifier is much higher than that of the CS amplifier. This is because in this case, the effect of the pole at $f_{p,\text{in}}$ on the overall frequency response of the amplifier becomes dominant. Since, owing to the Miller effect, C_{in} of the CS amplifier is much larger than that of the folded-cascode amplifier, its f_H is much lower.

To confirm this point, observe that C_{in} of the folded-cascode amplifier can be estimated by replacing R_L' in the equation used to compute C_{in} for the CS amplifier, with the total resistance R_{d1}, between the drain of Q_1 and ground. Here,

$$R_{d1} = r_{01} \| r_{03} \| R_{in2}$$

where R_{in2} is the input resistance of the common-gate transistor Q_2 and can be obtained using an approximation of the relationship found for input resistance of the common-gate amplifier:

$$R_{in2} \simeq \frac{r_{o2} + r_{o5}}{g_{m2} r_{o2}}$$

Thus,

$$R_{d1} \simeq r_{o1} \| r_{o3} \| \frac{r_{o2} + r_{o5}}{g_{m2} r_{o2}} = \frac{2}{g_{m2}}$$

Therefore, R_{d1} is much smaller than R_L' (in the CS amplifier $\simeq r_{o1} \| r_{o3}$). Hence, C_{in} of the designed folded-cascode amplifier is indeed much smaller than that of the CS amplifier because the $(1 + g_m R')$ multiplier is smaller for the folded-cascode device. This confirms that the folded-cascode amplifier is much less impacted by the Miller effect and, therefore, can achieve a much higher f_H when R_{sig} is large.

The midband gain of the folded-cascode amplifier can be significantly increased by replacing the current mirror Q_5–Q_6 with a current mirror having a larger output resistance, such as the cascode current mirror in Fig. 8.32, whose output resistance is approximately $g_m r_{o2}$. In this case, however, R_{in2} and, hence, R_{d1} increase, causing an increased Miller effect and a corresponding reduction in f_H.

Example MS.10.2 *continued*

Finally, it is interesting to observe that the frequency response of the folded-cascode amplifier shown in Fig. B.94 drops beyond f_H at approximately –20 dB/decade when $R_{sig} = 100\ \Omega$ and at approximately –40 dB/decade when $R_{sig} = 1\ \text{M}\Omega$. This is because when R_{sig} is small, the frequency response is dominated by the pole at $f_{p,\text{out}}$. However, when R_{sig} is increased, $f_{p,\text{in}}$ is moved closer to $f_{p,\text{out}}$, and both poles contribute to the gain falloff.

Example MS.11.1

A Two-Stage CMOS Op Amp with Series–Shunt Feedback

In this example, we will investigate the effect of applying a series–shunt feedback to the two-stage CMOS op amp whose schematic capture is shown in Fig. B.95.

The first stage is a differential pair Q_1–Q_2 (which is actively loaded with the current mirror formed by Q_3 and Q_4) with bias current supplied by a current mirror formed by Q_8 and Q_5, which utilizes the reference bias current I_{REF}. The second stage consists of Q_6, which is a common-drain amplifier actively loaded with a current source load (transistor Q_7).

For the implementation of this CMOS op amp, we will use a 0.18-μm CMOS technology for the MOSFETs and typical SPICE level-1 model parameters for this technology, including the intrinsic capacitance values. The targeted specifications are an unloaded dc open-loop voltage gain $|A_v| = 50$ V/V,

NMOS	
Vtn	0.5 V
|VAn|	12.5V
kn'	246.2uA/V^2
L	0.2 um
W	0.48 um

PMOS	
Vtp	–0.5 V
|VAp|	9V
kp'	–86.1 μA/V^2
L	0.2 um
W	0.64 um

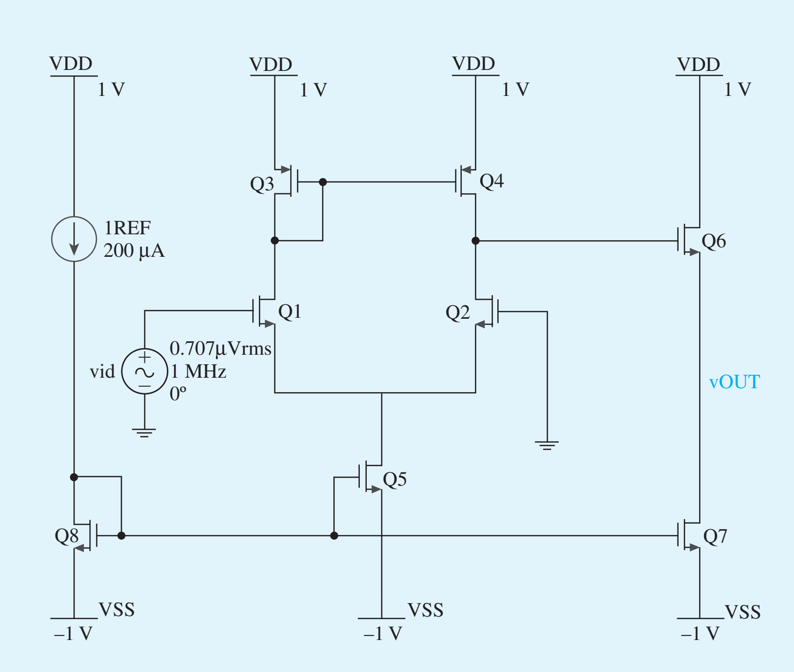


Figure B.95 Schematic capture of the two-stage CMOS op amp.

and closed-loop voltage gain $|A_f| = 10$ V/V, with each of transistors Q_1, Q_2, Q_3, and Q_4 biased at a drain current of 100 μA.

To achieve the targeted specifications, a biasing current $I_{REF} = 200$ μA is used, and the transistors Q_5, Q_6, Q_7, and Q_8 will be sized to conduct drain currents of 200 μA. The dc open-loop voltage gain for this amplifier is the product of the voltage gains of the two stages. Since the gain of the second stage (source follower) is approximately 1 V/V, the first stage must be designed to provide the full voltage gain of 50 V/V to achieve the specified open-loop voltage gain.

The amplifier specifications are summarized in Table B.25.

Table B.25 Two-Stage CMOS Op-Amp Specifications

Parameters	Value
$I_{(Q1,Q2,Q3,\text{and }Q4)}$	100 μA
$I_{(Q5,Q6,Q7,\text{and }Q8)}$	200 μA
$\|A_1\|\|$	50 V/V
$\|A_2\|$	1 V/V
$\|A_f\|$	10 V/V
V_{DD}	1 V
V_{SS}	−1 V

Hand Design

Design of the Two-Stage Op Amp The first stage of this CMOS op amp is identical to the first stage of the op amp we designed in Example MS.9.1, to which the reader is referred. Also, transistors Q_6 and Q_7 are sized to provide the bias current of 200 μA in the second stage.

As a result, using unit-size NMOS transistors with $W_n/L_n = 0.48\,\mu\text{m}/0.20\,\mu\text{m}$, and unit-size PMOS transistors with $W_p/L_p = 0.64\,\mu\text{m}/0.20\,\mu\text{m}$, the corresponding multiplicative factor m for each transistor can be calculated as found in Example MS.9.1 (with the difference here that Q_6 and Q_7 have the same dimensions). Table B.26 summarizes the relevant information and the calculated m values for the transistor.

Table B.26 Transistor Sizes

Transistor	I_D (μA)	m
1	100	6
2	100	6
3	100	14
4	100	14
5	200	13
6	200	13
7	200	13
8	200	13

Example MS.11.1 *continued*

Design of the Feedback Network First we need to determine the value of the feedback factor β for this series–shunt feedback amplifier. The β network can be implemented using a voltage divider, as shown in Fig. B.96. The resistor values are chosen large enough (in comparison to the output resistance of the designed two-stage op amp) to minimize the effect of loading. Therefore, effectively,

$$A \simeq A_v$$

where A is the open-loop gain of the amplifier (with loading). Now we can calculate the required feedback factor, β, as follows:

$$|A_f| = \frac{A_v}{1 + A_v\beta} = \frac{50}{1 + 50\beta} = 10\,\text{V/V}$$
$$\beta = 0.08$$

The resistor values of this voltage divider are selected to provide voltage divisions of 0.08 ($R_1 = 92$ kΩ and $R_2 = 8$ kΩ).

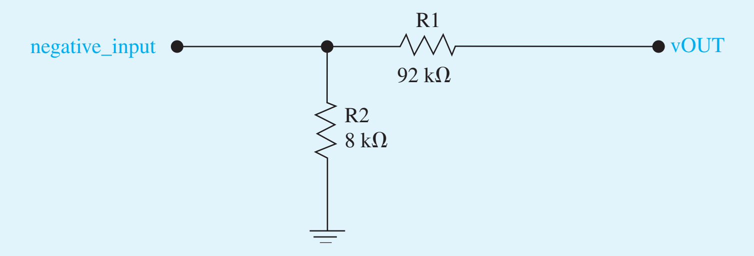


Figure B.96 β Network.

Simulation

Now we will simulate our designed circuit to verify our hand design and study the effect of feedback on the dc-gain, bandwidth, and output resistance of the amplifier.

Verifying A_v The schematic capture of the two-stage CMOS amplifier is in Fig. B.95. We can verify the dc voltage gain of this amplifier by performing frequency-response analysis as set up in Ch10_OpAmp_Ex_Av.ms10.

As can be seen from Fig. B.97, $|A_v| = 35.0$ dB $\simeq 56.2$ V/V, which is close to the targeted specification.

Verifying A The schematic capture of the A-circuit is given in Fig. B.98.

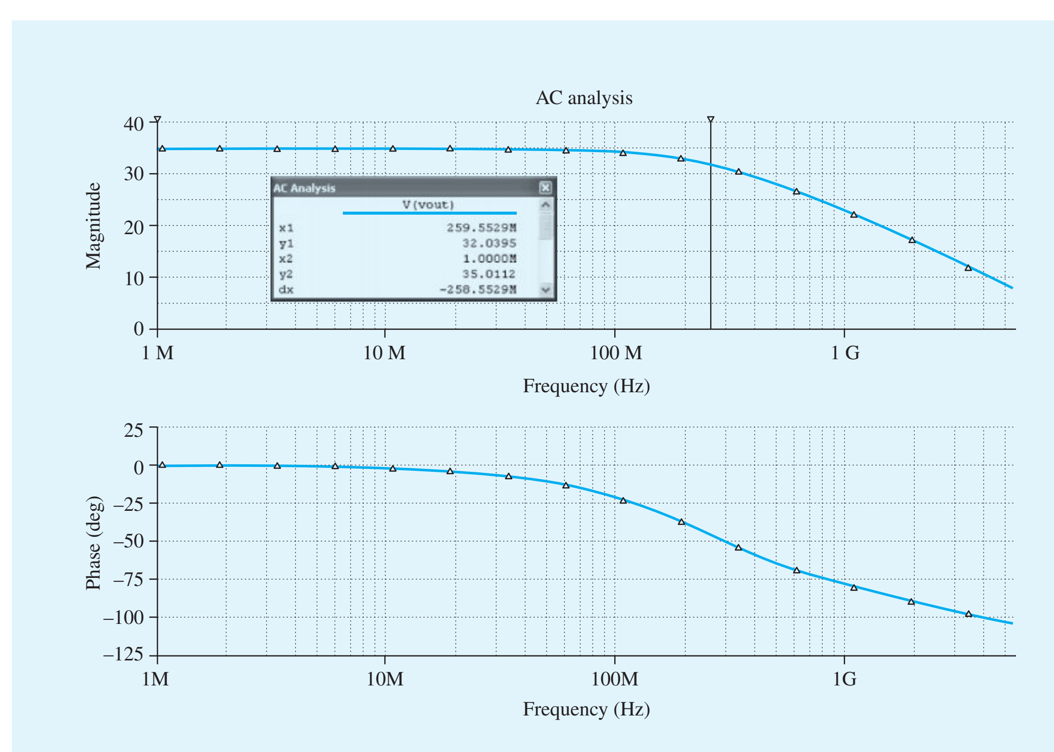
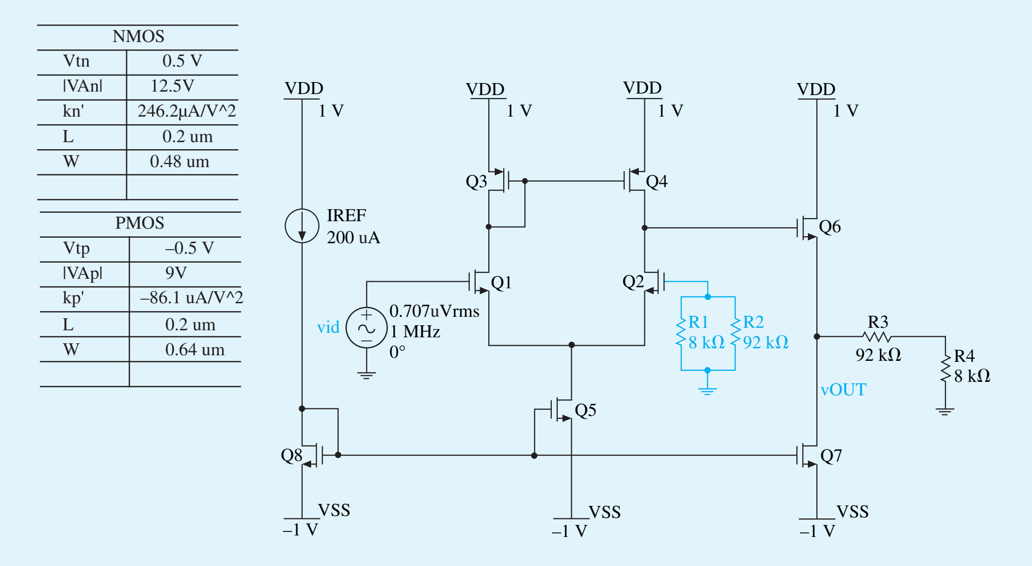


Figure B.97 Frequency response of the two-stage CMOS op-amp amplifier.



NMOS	
Vtn	0.5 V
|VAn|	12.5V
kn'	246.2μA/V^2
L	0.2 um
W	0.48 um

PMOS	
Vtp	–0.5 V
|VAp|	9V
kp'	–86.1 uA/V^2
L	0.2 um
W	0.64 um

Figure B.98 Schematic capture of the *A*-circuit.

We can verify the open-loop voltage gain of this circuit by performing a frequency-response analysis as set up in Ch10_OpAmp_Ex_A.ms10. As can be seen from Fig. B.99, $|A| = 34.9$ dB $\simeq 55.6$ V/V, which is close to the value of A_v. This supports our assumption of $A \simeq A_v$.



Example MS.11.1 *continued*

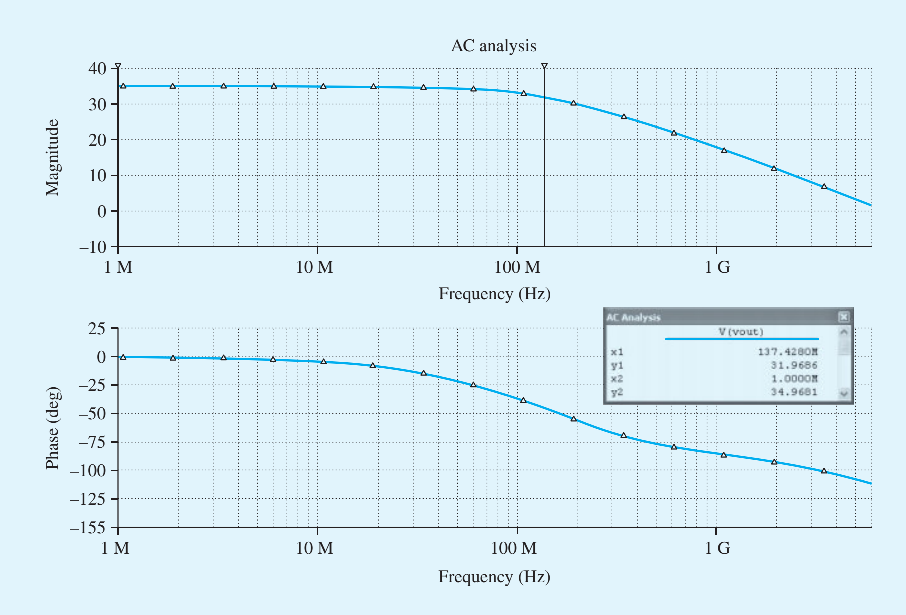


Figure B.99 Schematic capture of the *A*-circuit.

Verifying A_f The schematic capture of the closed-loop circuit is given in Fig. B.100. As can be seen from this schematic, the β-network establishes a series connection at the input and a shunt connection at the output of the original two-stage CMOS op amp.

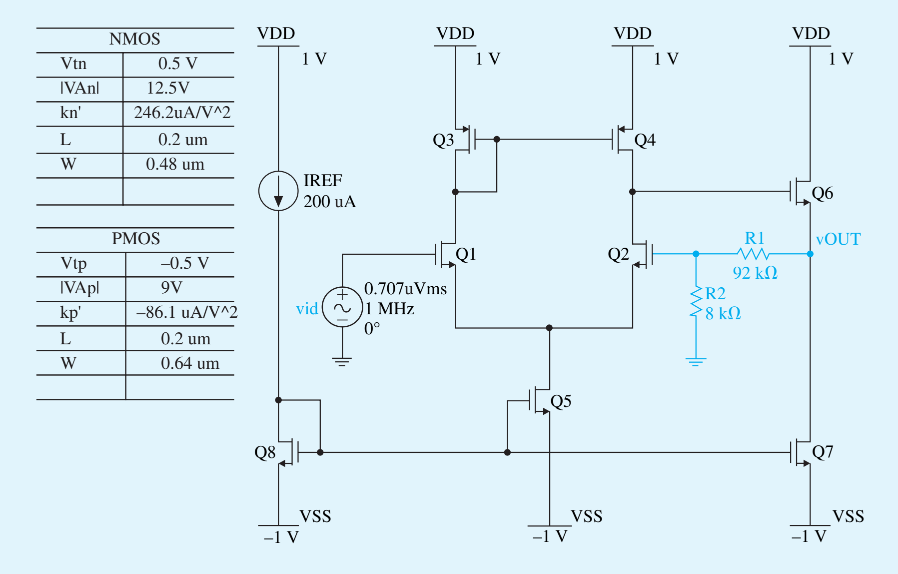


Figure B.100 Schematic capture of the closed-loop circuit.

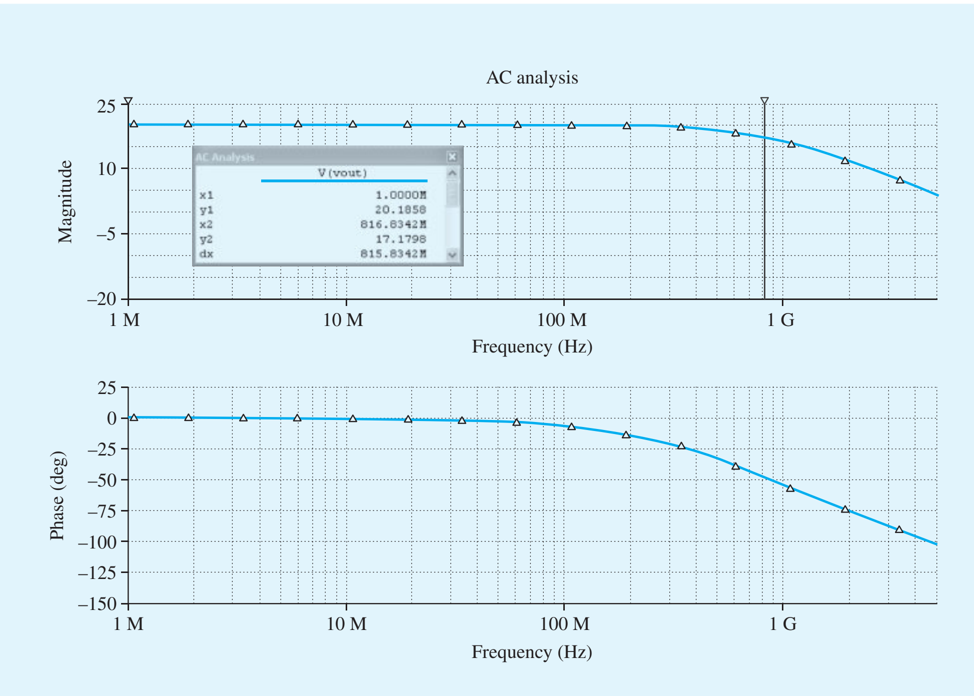


Figure B.101 Frequency response of the closed-loop circuit.

We can verify the closed-loop voltage gain by performing a frequency-response analysis as set up in Ch10_OpAmp_Ex_Af.ms10. As can be seen from Fig. B.101, $|A_f| = 20.2$ dB $\simeq 10.2$ V/V, which is close to the targeted specification for A_f.

Investigating the Effect of Feedback In addition to the frequency-response analysis, which provided information on the dc voltage gain and the 3-dB bandwidth, we used Multisim to find the output resistances of the open-loop and closed-loop circuits (as set up in Ch10_OpAmp_Ex_A.ms10 and Ch10_OpAmp_Ex_Af.ms10). Table B.27 summarizes our findings for open-loop (A-circuit) and closed-loop circuits.

Table B.27 Effect of Feedback on Gain, 3-dB Bandwidth, and Output Resistor

Circuit	Gain (V/V)	3-dB Bandwidth (MHz)	R_{out} (Ω)
Open loop	55.6	137	492.6
Closed loop	10.2	816	89.3

It can be seen from Table B.27 that the series–shunt feedback connection causes the dc voltage gain and the output resistance of the circuit to decrease by a factor of 5.5, while the 3-dB bandwidth increases by approximately the same factor. This factor is equal to $1 + A\beta$, the amount of the feedback. This is as expected and corresponds to what we learned in Chapter 11.

Example MS.12.1

Class B Bipolar Output Stage

In this example, we will design a class B output stage to deliver an average power of 20 W to an 8-Ω load. The schematic capture of a class B output stage implemented using BJTs is shown in Fig. B.102. We then will investigate various characteristics of the designed circuit such as crossover distortion and power-conversion efficiency. For this design, we are to select V_{CC} about 5 V greater than the peak output voltage in order to avoid transistor saturation and signal distortion.

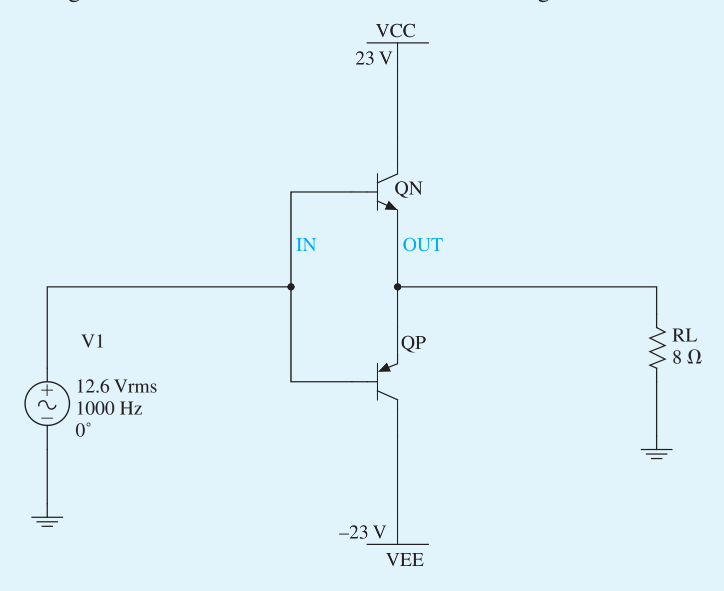


Figure B.102 Schematic capture of class B output stage.

The circuit specifications are summarized in Table B.28.

Table B.28 Class B Output Stage Specifications

Specification	Value
P_L	20 W
R_L	8 Ω
V_{CC}	$\hat{V} + 5$ V

Hand Design

We know from Eq. (12.12) that

$$P_L = \frac{1}{2}\frac{\hat{V}^2}{R_L}$$

Example MS.12.1 *continued*

Thus

$$V = \sqrt{2P_L R_L} = \sqrt{2 \times 20 \times 8}$$
$$V = 17.9 \text{ V}$$

which leads to $V_{CC} = 23$ V.

The peak current drawn from the supply will be

$$I_O = \frac{\hat{V}}{R_L} = \frac{17.9}{8}$$
$$I_o = 2.24 \text{ A}$$

Now we can use Eq. (12.13) to calculate the average power drawn from each of the supplies

$$P_{S+} = P_{S-} = \frac{1}{\pi}\frac{\hat{V}}{R_L}V_{CC} = \frac{1}{\pi} \times 2.24 \times 23$$
$$P_{S+} = P_{S-} = 16.4 \text{ W}$$
$$P_S = P_{S+} + P_{S-} = 16.4 + 16.4 = 32.8 \text{ W}$$

Therefore, the power-conversion efficiency, η, is

$$\eta = \frac{P_L}{P_S} \times 100\% = \frac{17.9}{32.8} \times 100\% = 61\%$$

Now we can utilize Eq. (12.22) to calculate the maximum power dissipated in each of the transistors as

$$P_{DN\max} = P_{DP\max} = \frac{V_{CC}^2}{\pi^2 R_L} = \frac{(23)^2}{\pi^2 \times 8}$$
$$P_{DN\max} = P_{DP\max} = 6.7 \text{ W}$$

Simulation

Next, we use Multisim to verify the operation of the class B output stage designed above. For simulation purposes, we will use discrete-power transistors MJE243 and MJE253 (from ON Semiconductor), which are rated for a maximum continuous collector current $I_{C\max} = 4$ A and a maximum collector–emitter voltage $V_{CE\max}$ of 100 V.

Load Power P_L To measure the amount of power delivered to the load, we will utilize Transient Analysis in Multisim as set up in Ch11_Class_B_Ex.ms10. The transient analysis simulation is performed over the interval 0 ms to 2 ms, and the waveforms of the voltage at the output node and the output current are plotted in Fig. B.103.

As can be seen in Fig. B.103, the peak voltage amplitude is approximately 16.9 V and the peak current amplitude is 2.1 A. Upon a closer look at the current and voltage waveforms, we can observe that both exhibit crossover distortion. The bottom graph in Fig. B.103 shows the instantaneous and the average

Example MS.12.1 *continued*

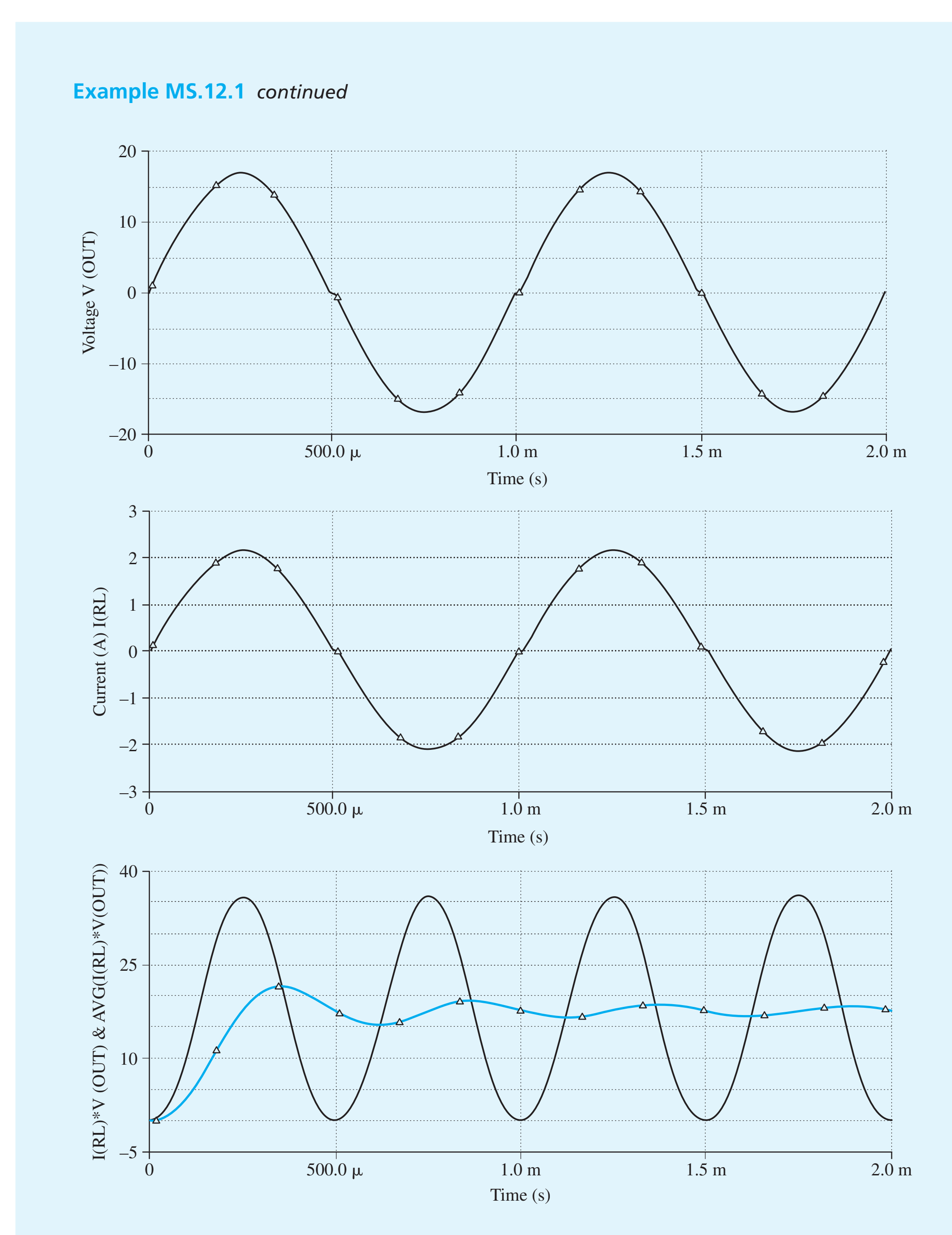


Figure B.103 Load voltage, current, and instantaneous and average load power.

power dissipated in the load resistance. These waveforms were obtained by multiplying the current and voltage waveforms, and by taking the running average for the average power, P_L. The transient behavior of average power eventually settles into a quasi-constant steady-state value of 17.6 W.

Supply Power, P_S Similarly, we can plot instantaneous voltage and current at the V_{CC} and V_{EE} nodes to measure the value of P_S. Figure B.104 shows the voltage, current, instantaneous, and average power for $+V_{CC}$. We can plot these quantities for $-V_{EE}$ as well. However, owing to symmetry, we do not need to generate plots for the negative voltage supply. The average power provided by $+V_{CC}$, P_{S+}, is 15 W. Therefore, the total power provided by both voltage supplies is 30 W.

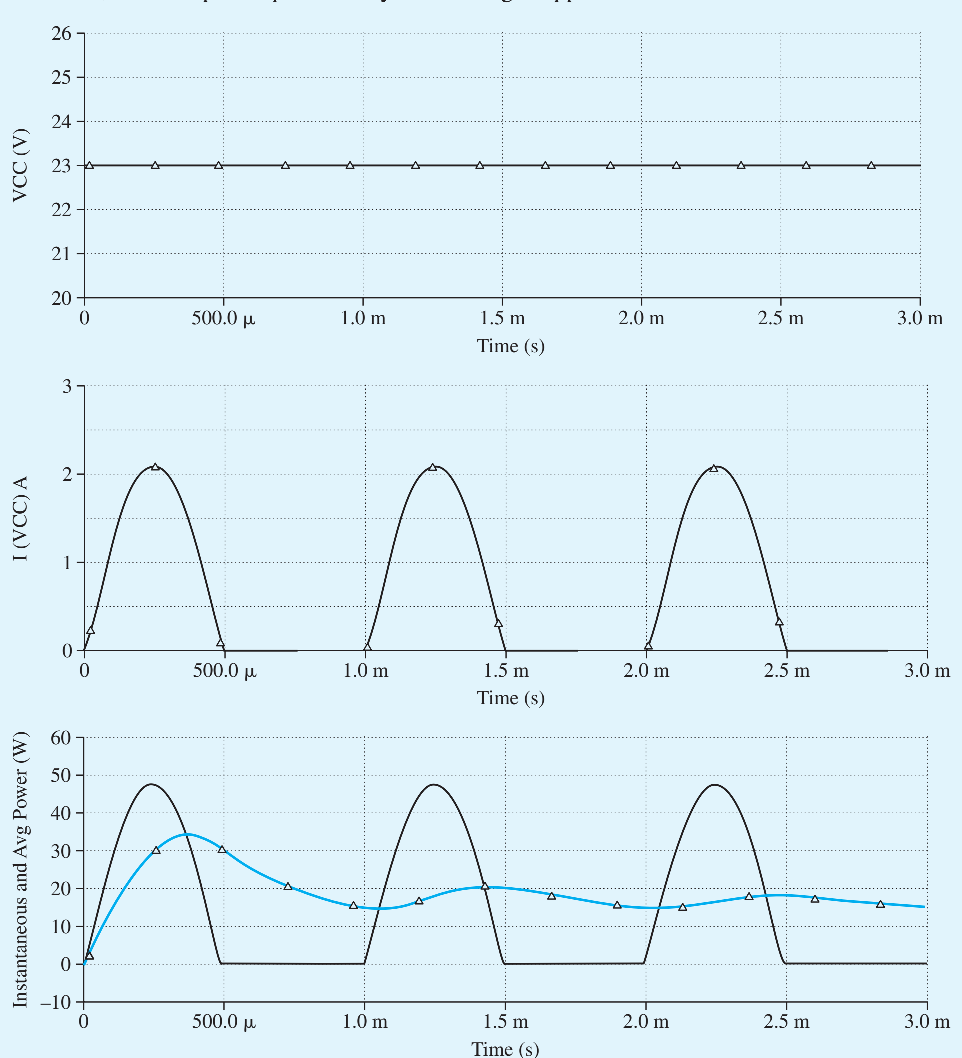


Figure B.104 Supply voltage, current, and instantaneous and average supply power.

Power-Conversion Efficiency, η Now we can calculate the power-conversion efficiency of the simulated circuit as follows.

$$\eta = \frac{P_L}{P_S} \times 100\% = \frac{17.6}{30} \times 100\% = 58.6\%$$

Example MS.12.1 *continued*

Transistor Power Dissipation, P_D Figure B.103 shows voltage, current, instantaneous and average power plots for Q_p only. A similar plot can be obtained for Q_n to measure the power dissipated in the *npn* device. As expected, the voltage waveform is a sinusoid, and the current waveform consists of half-sinusoids. The waveform of instantaneous power is rather unusual. It indicates the presence of some distortion as a result of driving the transistors rather hard. This can be verified by reducing the amplitude of the input signal. Students are encouraged to investigate this point. The average power dissipated in Q_p, as measured from Fig. B.105, is approximately 6 W. Therefore, the total power dissipated in the transistors is 12 W.

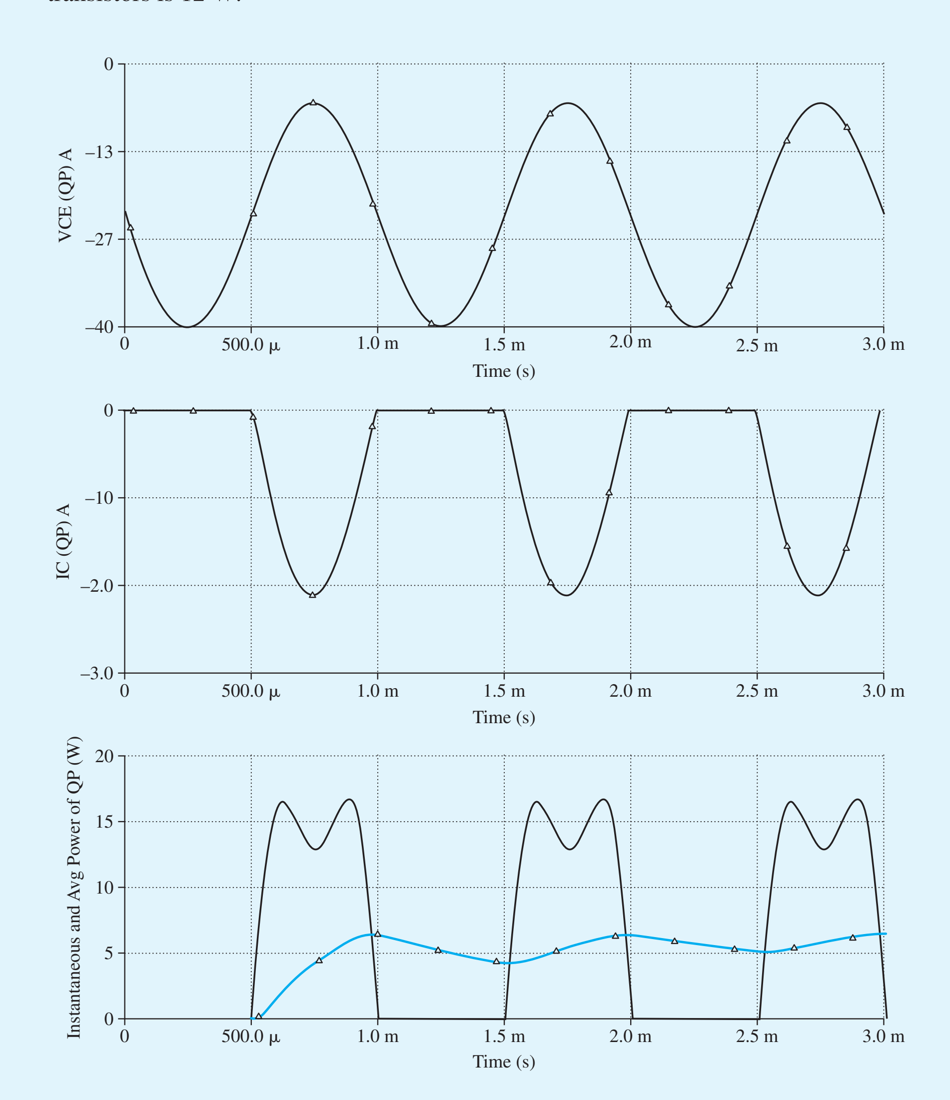


Figure B.105 Voltage, current, and instantaneous and average power for Q_1.

The simulation results and hand-design calculations are summarized in Table B.29. Observe that the values are quite close, which verifies our design of the class B output stage.

Table B.29 Summary of Simulation and Hand-Design Results

Measurement	Hand Design	Simulated	Error %
P_L	17.9 W	17.6 W	1.7
P_S	32.8 W	30 W	8.5
P_D	13.4 W	12 W	10.4
n	61%	58.6%	3.9

Crossover Distortion We can further investigate the crossover distortion of this circuit by utilizing the voltage transfer characteristics (VTC) curve of the class B output stage. The plot is obtained through a dc sweep analysis in Multisim where v_{IN} is swept over the range –10 V to 10 V in 1.0-mV increments. From the resulting VTC curve, shown in Fig. B.106, we can see that the dead band extends from –0.605 V to 0.56 V. The effect of crossover distortion can be quantified by performing a Fourier analysis on the output voltage in Multisim.

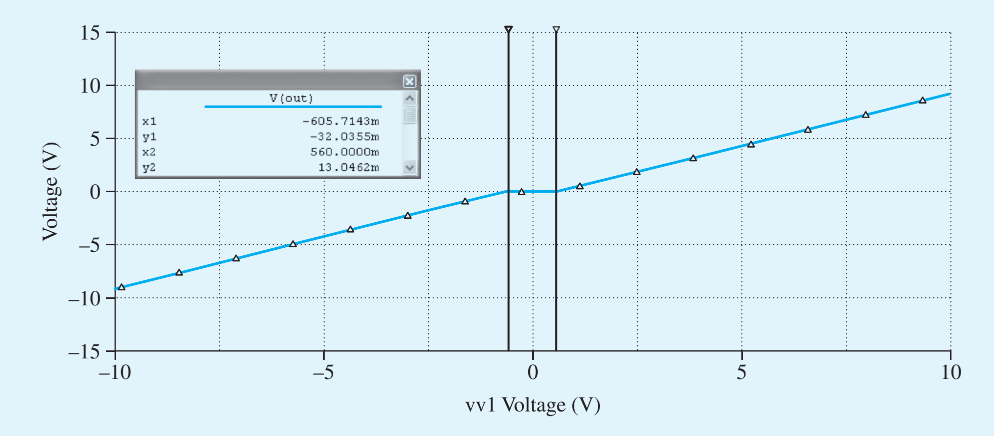


Figure B.106 VTC of class B output stage.

Total Harmonic Distortion (THD) This analysis decomposes the waveform generated via transient analysis into its Fourier-series components. Furthermore, Multisim computes the THD of the output waveform, and the results are shown in Fig. B.107.

Example MS.12.1 *continued*

4	THD:	2.12781 %				
5	Gridsize:	256				
6	Interpolation Degree:	1				
7						
8	Harmonic	Frequency	Magnitude	Phase	Norm. Mag	Norm. Phase
9	1	1000	16.7351	-0.0017584	1	0
10	2	2000	0.0104931	90.328	0.00062701	90.3297
11	3	3000	0.274201	-179.99	0.0163847	-179.98
12	4	4000	0.00492676	91.2479	0.000294396	91.2497
13	5	5000	0.172473	179.997	0.0103061	179.999
14	6	6000	0.00165276	92.6225	9.87598e-005	92.6243
15	7	7000	0.117776	179.948	0.00703764	179.949
16	8	8000	0.00148745	95.535	8.8882e-005	95.5368

Figure B.107 Fourier-series components of the output voltage and class B output stage THD.

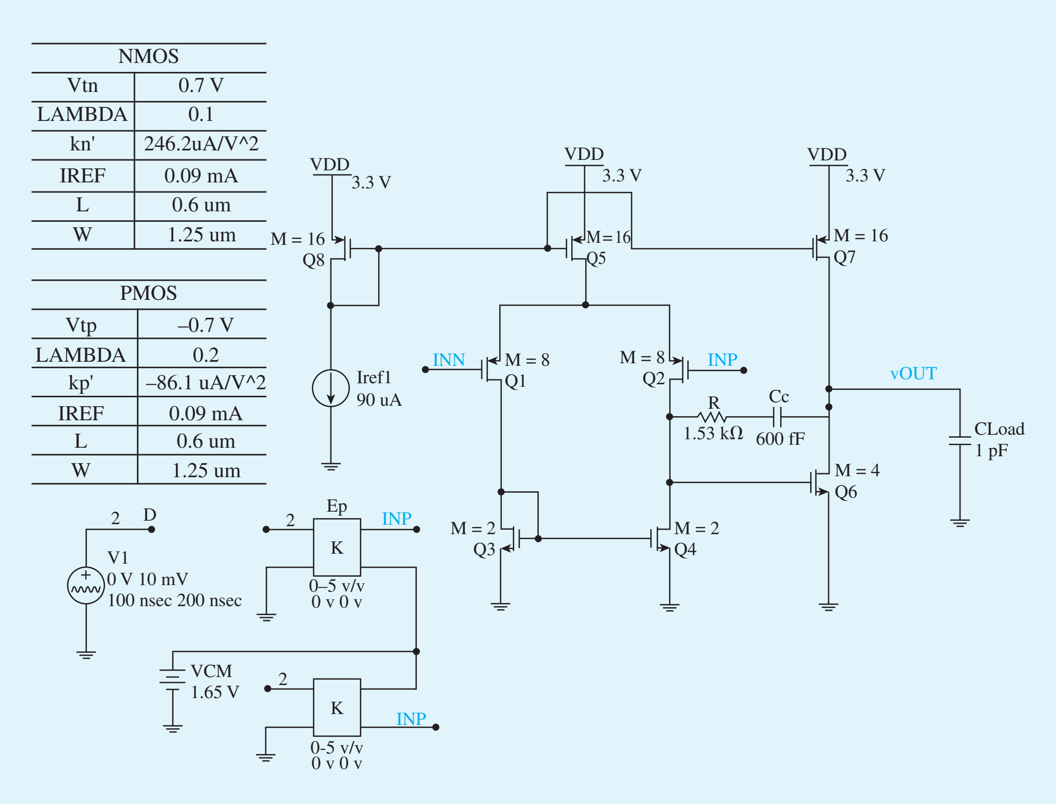
From the Fourier analysis, we note that the waveform is rather rich in odd harmonics and that the resulting THD is 2.13%, which is rather high.

Example MS.13.1

A Two-Stage CMOS Op Amp with Frequency Compensation

In this example, we will use Multisim to aid in designing the frequency compensation of the two-stage CMOS circuit whose schematic is shown in Fig. B.108. Multisim will then be employed to determine the frequency response and the slew rate of the op amp. We will assume a 0.5-μm CMOS technology for the MOSFETs and use typical SPICE level-1 model parameters for this technology.

The op-amp circuit in Fig. B.108 is designed using a reference current $I_{REF} = 90\ \mu A$, a supply voltage $V_{DD} = 3.3$ V, and a load capacitor $C_L = 1$ pF. Unit-size transistors with $W/L = 1.25\,\mu m/0.6\,\mu m$ are used for both the NMOS and PMOS devices. The transistors are sized for an overdrive voltage $V_{OV} = 0.3$ V. The corresponding multiplicative factors are shown in Fig. B.108.



NMOS	
Vtn	0.7 V
LAMBDA	0.1
kn'	246.2uA/V^2
IREF	0.09 mA
L	0.6 um
W	1.25 um

PMOS	
Vtp	–0.7 V
LAMBDA	0.2
kp'	–86.1 uA/V^2
IREF	0.09 mA
L	0.6 um
W	1.25 um

Figure B.108 Schematic capture of the two-stage CMOS op amp.

In Multisim, the common-mode input voltage V_{CM} of the op-amp circuit is set to $V_{DD}/2 = 1.65$ V, and dc operating-point analysis is performed to determine the dc bias conditions. Using the values found from the simulation results for the small-signal parameters of the MOSFETs, we obtain

$$G_{m1} = 0.333 \text{ mA/V}$$
$$G_{m2} = 0.650 \text{ mA/V}$$
$$C_1 = 26.5 \text{ fF}$$
$$C_2 = 1.04 \text{ pF}$$

using Eqs. (13.7), (13.14), (13.25), and (13.26) respectively. Recall that G_{m1} and G_{m2} are the transconductances of, respectively, the first and second stages of the op amp. Capacitors C_1 and C_2 represent the total capacitance to ground at the output nodes of, respectively, the first and second stage of the op amp.

Then, using Eq. (13.35), the frequency of the second, nondominant, pole can be found as

$$f_{P2} \simeq \frac{G_{m2}}{2\pi C_2} = 97.2\,\text{MHz}$$

Example MS.13.1 *continued*

To place the transmission zero, given by Eq. (13.31), at infinite frequency, we select

$$R = \frac{1}{G_{m2}} = 1.53\,\text{k}\Omega$$

Now, using Eq. (13.42), the phase margin of the op amp can be expressed as

$$PM = 90° - \tan^{-1}\left(\frac{f_t}{f_{P2}}\right)$$

where f_t is the unity-gain frequency, given in Eq. (13.83):

$$f_1 = \frac{G_{m1}}{2\pi C_C}$$

Using the above two equations we determine that compensation capacitors of $C_C = 0.78$ pF and $C_C = 2$ pF are required to achieve phase margins of $PM = 55°$ and $PM = 75°$, respectively.

Next, an ac-analysis simulation is performed in Multisim to compute the frequency response of the op amp and to verify the foregoing design values (as set up in Ch12_Two_Stage_CMOS_OpAmp_Ex_Freq-Resp.ms10). It was found that, with $R = 1.53$ kΩ, we needed $C_C = 0.6$ pF and $C_C = 1.8$ pF to set $PM = 55°$ and $PM = 75°$, respectively. We note that these values are reasonably close to those predicted by hand analysis. The corresponding frequency responses for the compensated op amps are plotted in Figs. B.109 and B.110. For comparison, we also show the frequency response

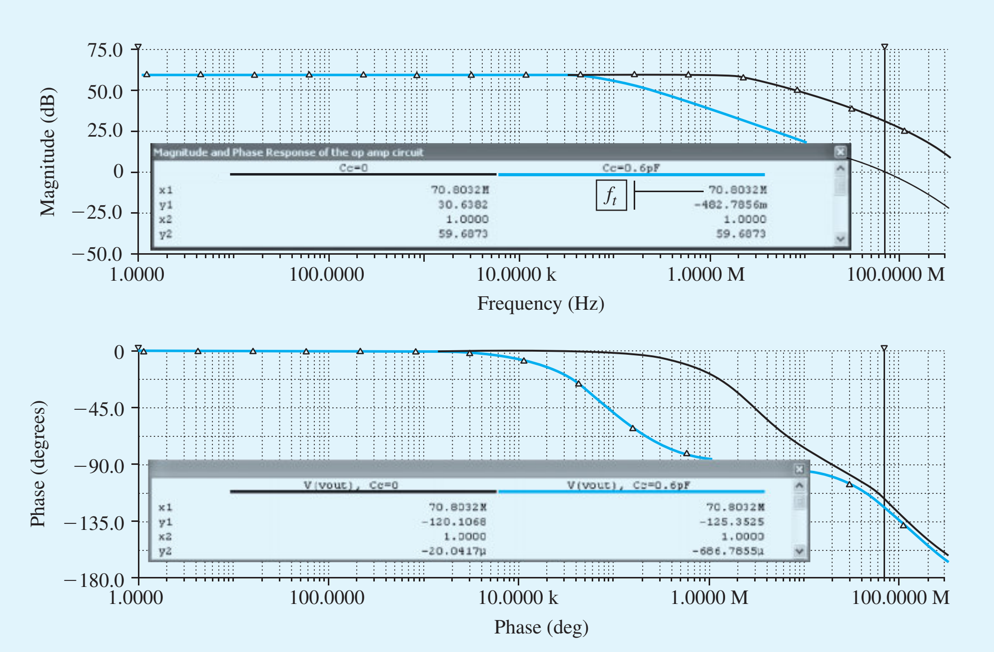


Figure B.109 Magnitude and phase response of the op-amp circuit with $R = 1.53$ k, $C_C = 0$ (no frequency compensation), and $C_C = 1.8$ pF ($PM = 75°$).

of the uncompensated op amp ($C_C = 0$). Observe that the unity-gain frequency f_t drops from 70.8 MHz to 26.4 MHz as C_c is increased to improve *PM*.

Rather than increasing the compensation capacitor C_C to improve the phase margin, the value of the series resistor *R* can be increased: for a given C_C, increasing *R* above $1/G_{m2}$ places the transmission zero at a negative real-axis location (Eq. 13.31), where the phase it introduces *adds* to the phase margin. Thus, *PM* can be improved without affecting f_t. To verify this point, we set C_C to 0.6 pF and simulate the op-amp circuit in Multisim for the cases of $R = 1.53$ kΩ and $R = 3.2$ kΩ. The corresponding frequency response is plotted in Fig. B.111. Observe how f_t is approximately independent of *R*. However, by increasing *R*, we can improve *PM* from 55° to 75°.

Increasing the *PM* is desirable because it reduces the overshoot in the step response of the op amp. To verify this point, we simulate in Multisim the step response of the op amp for $PM = 55°$ and $PM = 75°$. To do that, we connect the op amp in a unity-gain configuration, apply a small (10-mV) pulse signal at the input with very short (1-ps) rise and fall times to emulate a step input, perform a transient analysis simulation (as set up in Ch12_Two_Stage_CMOS_OpAmp_Ex_Small-Signal.ms10), and plot the output voltage as shown in Fig. B.112. Observe that the overshoot in the step response drops from 15% to 1.4% when the phase margin is increased from 55° to 75°.

We conclude this example by computing *SR*, the slew rate of the op amp. From Eq. (13.45), we have

$$SR = 2\pi f_t V_{OV} = \frac{G_{m1}}{C_C} V_{OV} = 166.5\,\text{V}/\mu\text{s}$$

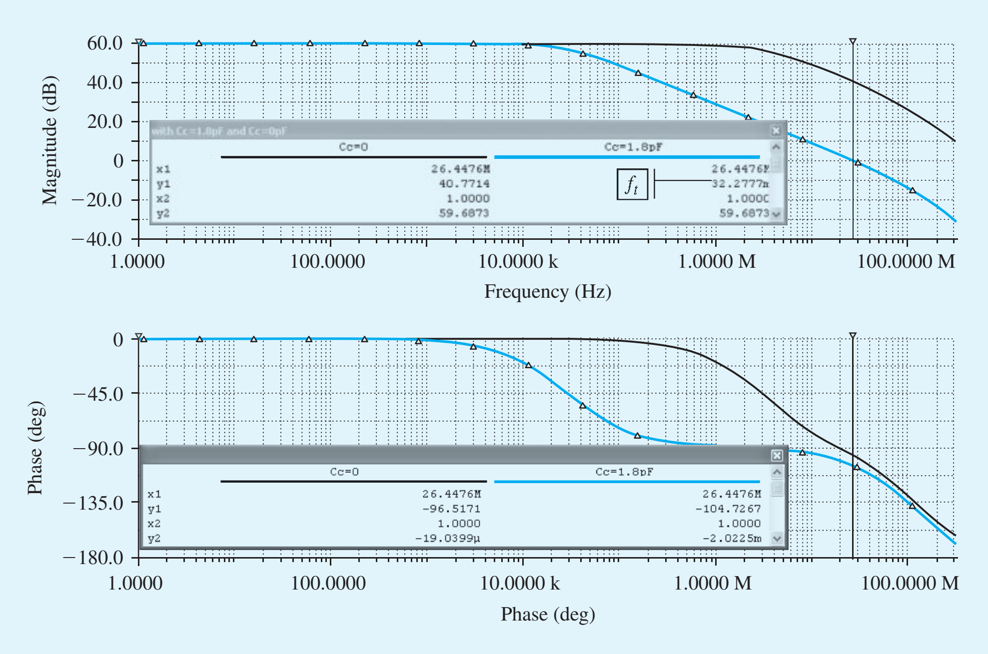


Figure B.110 Magnitude and phase response of the op-amp circuit with $R = 1.53$ kΩ, $C_C = 0$ (no frequency compensation), and $C_C = 1.8$ pF ($PM = 75°$).

Example MS.13.1 *continued*

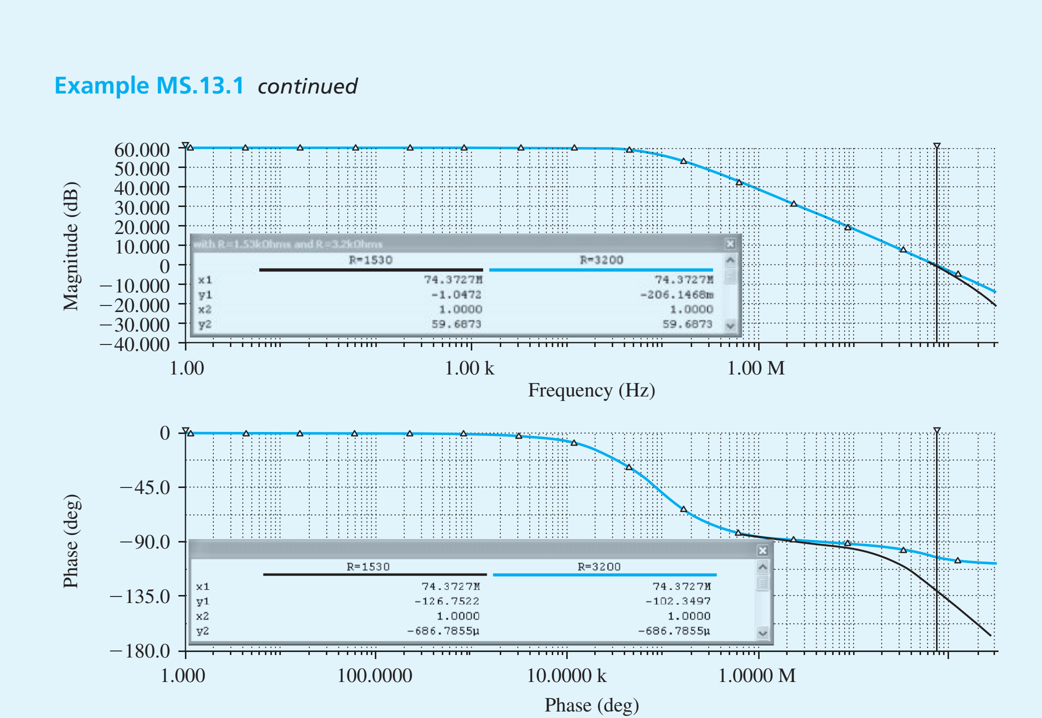


Figure B.111 Magnitude and phase response of the op-amp circuit with $C_C = 0.6$ pF, $R = 1.53$ kΩ ($PM = 55°$), and $R = 3.2$ kΩ ($PM = 75°$).

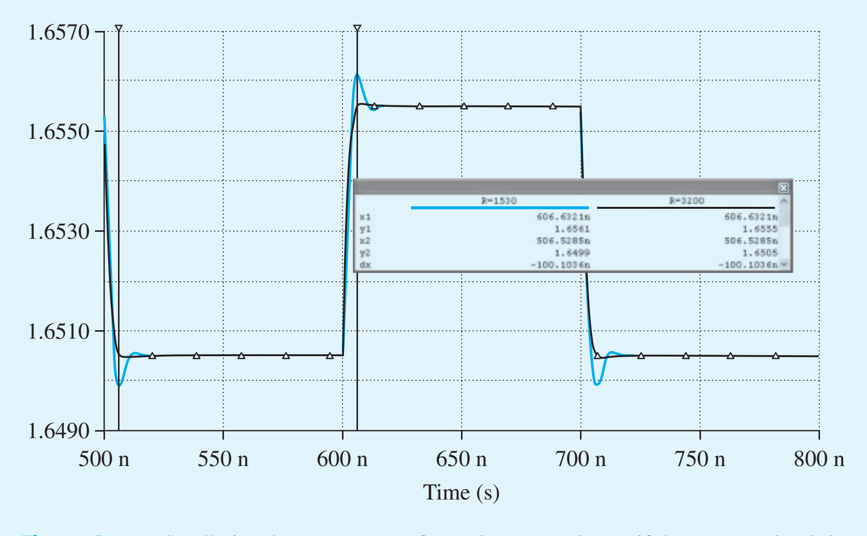


Figure B.112 Small-signal step response (for a 10-mV step input) if the op-amp circuit is connected in a unity-gain configuration: $PM = 55°$ ($C_C = 0.6$ pF, $R = 1.53$ kΩ) and $PM = 75°$ ($C_C = 0.6$ pF, $R = 3.2$ kΩ).

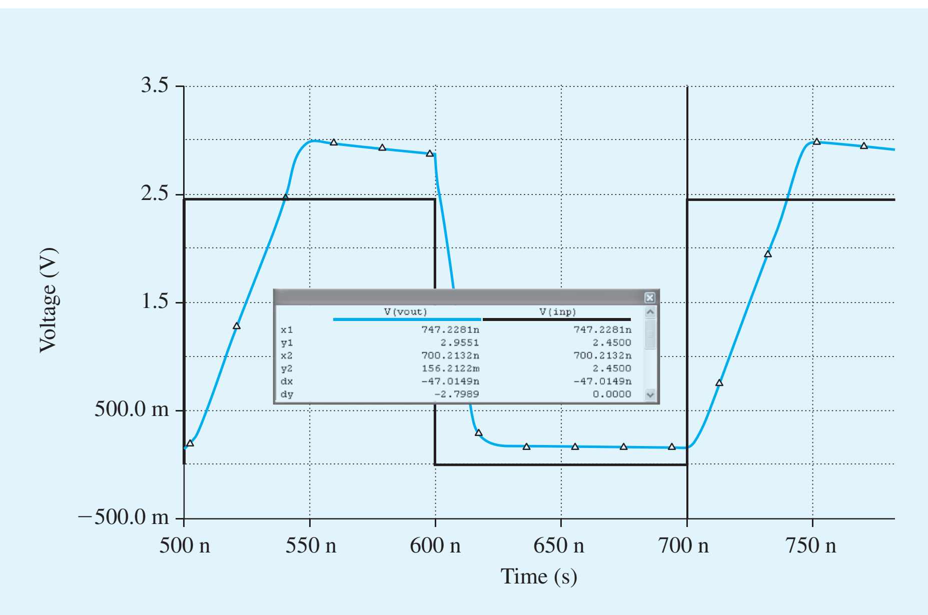


Figure B.113 Large-signal step response (for a 2.2-V step input) if the op-amp circuit is connected in a unity-gain configuration. The slope of the rising and falling edges of the output waveform correspond to the slew rate of the op amp.

when $C_C = 0.6$ pF. Next, to determine SR using Multisim, we again connect the op amp in a unity-gain configuration and perform a transient analysis simulation (as set up in Ch12_Two_Stage_CMOS_OpAmp_Ex_Large-Signal.ms10). However, we now apply a large pulse signal (2.2 V) at the input to cause slew-rate limiting at the output. The corresponding output voltage waveform is plotted in Fig. B.113. The slope of the slew-rate-limited output waveform corresponds to the slew rate of the op amp and is found to be $SR = 160$ V/μs and 60 V/μs for the negative- and positive-going output, respectively. These results, with the unequal values of SR in the two directions, differ from those predicted by the simple model for the slew-rate limiting of the two-stage op-amp circuit. The difference can perhaps be said to be a result of transistor Q_4 entering the triode region and its output current (which is sourced through C_C) being correspondingly reduced. Of course, the availability of Multisim should enable the reader to explore this point further.

Example MS.14.1

The CMOS Inverter

In this example, we will use Multisim to design a CMOS inverter whose schematic capture is shown in Fig. B.114. We will assume a 0.18-μm CMOS technology for the MOSFETs and use typical SPICE level-1 model parameters for this technology, including the intrinsic capacitance values. This model does not take into account the short-channel effects for this technology. Also, the load capacitance is assumed

Example MS.14.1 *continued*

to be dominated by the extrinsic component C_{ext} (resulting from the wiring and the input capacitance of the driven gates), where the value used in this example is 15 fF. We will begin with an approximate hand-analysis design. We will then use Multisim to verify that the designed circuit meets the specifications. The targeted specification for this inverter is a high-to-low propagation delay (t_{PHL}) of less than 45 ps. Once designed, the other characteristics of this inverter such as low-to-high propagation delay (t_{PLH}), noise margins, and threshold voltage will be investigated.

The inverter specifications are summarized in Table B.30.

Table B.30 CMOS Inverter Specifications

Parameters	Value
t_{PHL}	45 ps
C_L	15 fF
V_{DD}	1.8 V

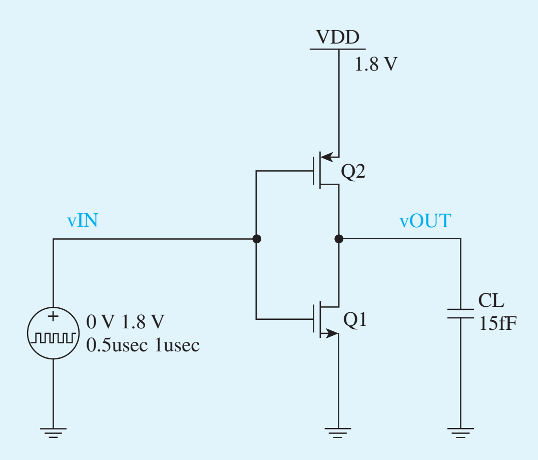


Figure B.114 Schematic capture of the CMOS inverter.

Hand Design

For the design of this inverter we choose $L = 0.2\ \mu m$, so we have $L_{eff} = 0.180\ \mu m$. As mentioned earlier, to minimize area, all channels are usually made equal to the minimum length permitted by the given technology. To meet the specified t_{PHL}, we need to size $(W/L_{eff})_n$ carefully. Once sized, $(W/L_{eff})_p = 2(W/L_{eff})_n$ is chosen, which is a compromise between area, noise margins, and t_{PLH}.

The value of t_{PHL} can be estimated using Eq. (14.46) as

$$t_{PHL} = \frac{\alpha_n C}{k_n'\left(\dfrac{W}{L_{eff}}\right)V_{DD}} = \frac{15 \times 10^{-15}\alpha_n}{246.2 \times 10^{-6}\left(\dfrac{W}{L_{eff}}\right)1.8}$$

where α_n is a factor determined by the relative values of V_t and V_{DD} $(V_{tn}/V_{DD=0.5/1.8=0.278})$:

$$\alpha_n = \frac{2}{7/4^{-3}\left(V_{tn}/V_{DD}\right) + \left(V_{tn}/V_{DD}\right)^2} = 2.01$$

Based on the above equations, the specified t_{PHL} can be achieved by selecting the ratio $\left(W/L_{\text{eff}}\right)_n = 1.5$ and consequently $\left(W/L_{\text{eff}}\right)_n = 3$. Table B.31 summarizes the relevant sizing information for each transistor. The third column of this table shows the transconductance parameter values for each transistor (which are typical values for 0.18-μm CMOS technology).

Table B.31 Transistor Sizes

Transistor	W (μm)	L_{eff} (μm)	k' (μA/V^2)
NMOS	0.27	0.18	246.2
PMOS	0.54	0.18	86.1

Note that for the selected width values, the intrinsic capacitances C_{gd1} and C_{gd2} are insignificant in comparison to the load capacitance. This confirms our initial assumption that in our hand calculations of delay, we could neglect C_{gd1} an C_{gd2} (which vary proportionally with width).

Simulation

Verifying Propagation Delay To investigate the dynamic operation of the inverter and to verify that the design meets the specified t_{PHL}, we apply an ideal pulse signal at the input and perform a transient analysis, as set up in Ch13_CMOS_Inverter_tPHL_Ex.ms10. Then, we plot the input and output waveforms as shown in Fig. B.115. Based on the simulated response, $t_{PHL} = 40.5$ ps (as indicated in Fig. B.115). Similarly, we obtain $t_{PLH} = 60.3$ ps, resulting in the inverter propagation delay (t_p) of 50.4 ps. Therefore, the specified high-to-low propagation delay specification is met, and t_p takes a reasonable value.

Voltage Transfer Characteristic (VTC) To compute both the VTC of the inverter and its supply current at various values of the input voltage V_{in}, we apply a dc voltage source at the input and perform a dc sweep with V_{in} swept over the range 0 to V_{DD}, as set up in Ch13_CMOS_Inverter_VTC_Ex.ms10. The resulting VTC is plotted in Fig. B.116. Note that the slope of the VTC in the switching region (where the NMOS and PMOS devices are both in saturation) is not infinite as predicted from the simple theory presented earlier. Rather, the nonzero value of λ causes the inverter gain to be finite. The two points on the VTC at which the inverter gain is unity (i.e., the VTC slope is –1 V/V) and that determine V_{IL} and V_{IH} are indicated in Fig. B.116. The corresponding noise margins are $NM_L = 0.76$ V and $NM_H = 0.81$ V. Note that the design provides high tolerance to noise, since noise margins are reasonably high (NM_L and NM_H are 42% and 45% of the supply voltage). This implies that the inverter would provide the correct logic output for an input noise variation of up to approximately 40% of the V_{DD}.



Example MS.14.1 *continued*

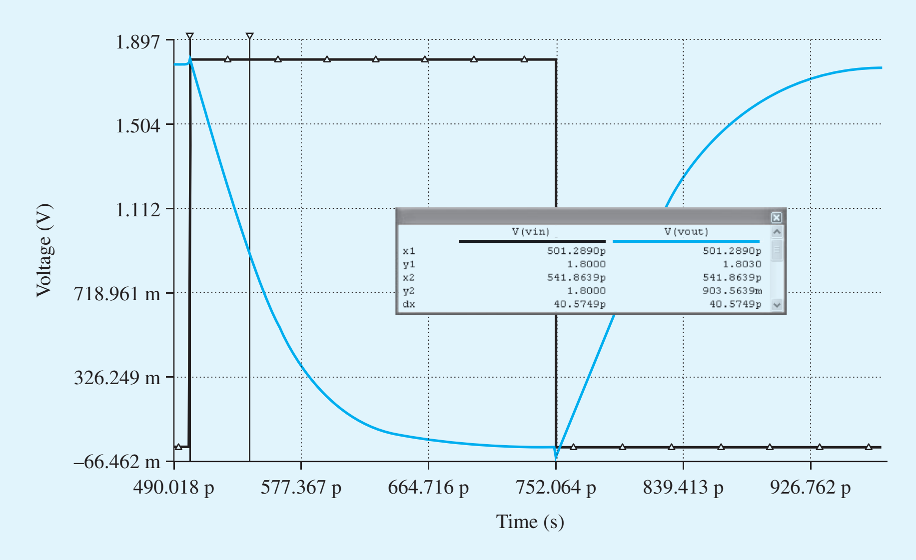


Figure B.115 Time-domain response of the CMOS inverter to measure t_{PHL}.

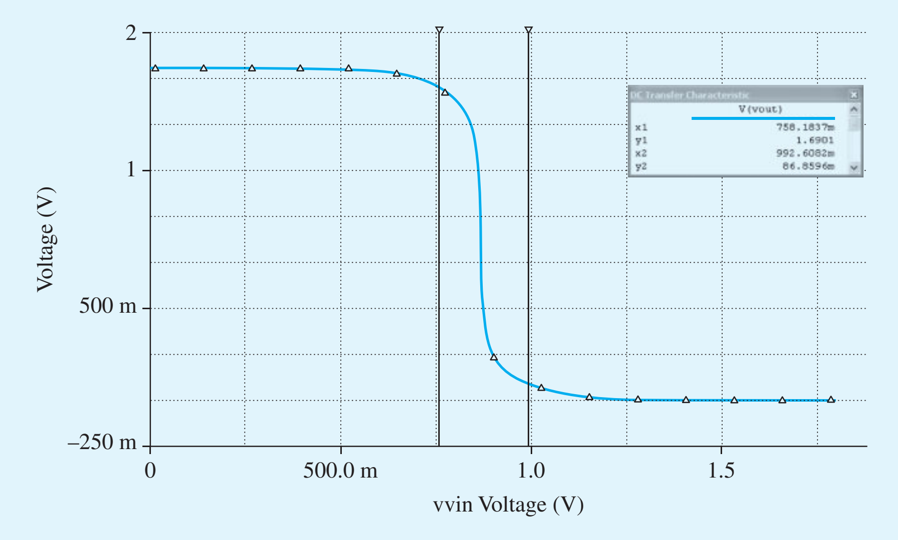


Figure B.116 Output voltage versus input voltage for the inverter (to measure low and high noise margins N_{ML} and N_{MH}).

The threshold voltage V_M of the CMOS inverter is defined as the input voltage v_{IN} that results in an identical switching output voltage v_{OUT}, that is,

$$V_M = v_{IN}\big|v_{OUT} = v_{IN}$$

Thus, as shown in Fig. B.117, V_M is at the intersection of the VTC with the straight line corresponding to $v_{OUT} = v_{IN}$. This line can be simply generated by plotting v_{IN} on the vertical axis, in addition to v_{OUT}. Note that $V_M = 0.87$ V, which is very close to the desired value of $V_{DD}/2 = 0.9$ V, as desired.

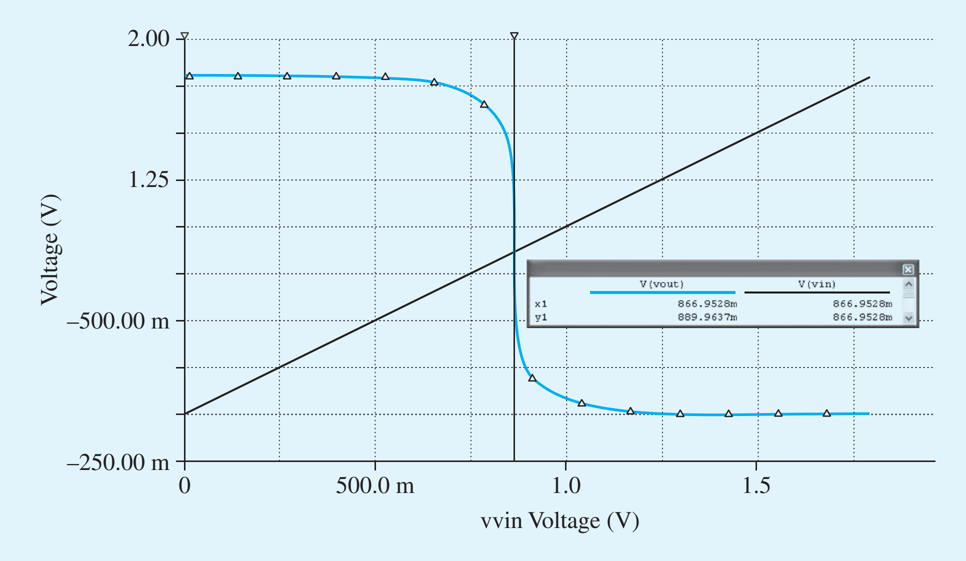


Figure B.117 Output voltage versus input voltage for the inverter (to the threshold voltage measure V_t).

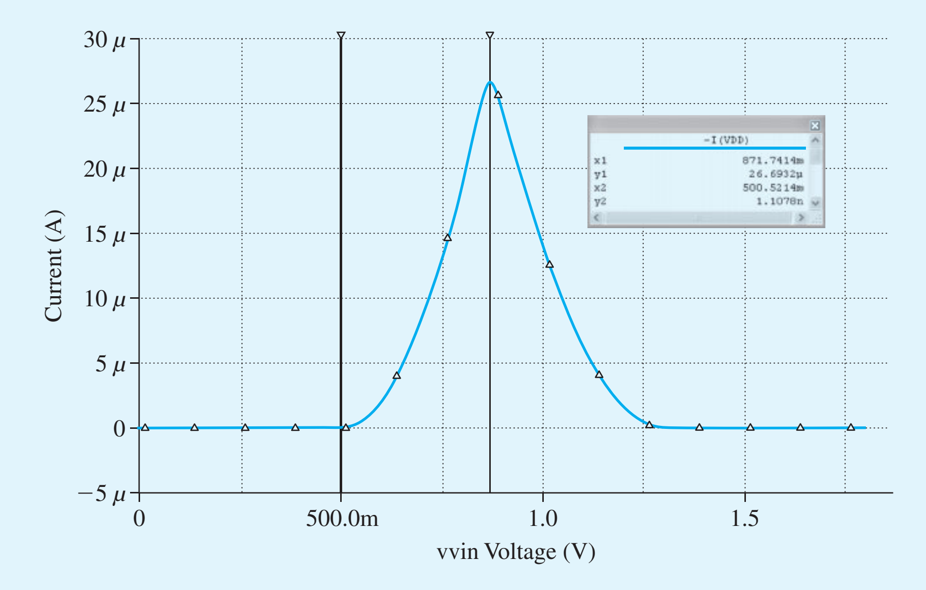


Figure B.118 Supply current versus input voltage for the inverter.

Finally, the supply current is plotted versus input voltage in Fig. B.118. Observe that in the transition region, where the inverter is switching, the current is no longer zero. Specifically, the peak current occurs at the inverter threshold voltage.

APPENDIX C

TWO-PORT NETWORK PARAMETERS

Introduction

At various points throughout the text, we make use of some of the different possible ways to characterize linear two-port networks. A summary of this topic is presented in this appendix.

C.1 Characterization of Linear Two-Port Networks

A two-port network (Fig. C.1) has four port variables: V_1, I_1, V_2, and I_2. If the two-port network is linear, we can use two of the variables as excitation variables and the other two as response variables. For instance, the network can be excited by a voltage V_1 at port 1 and a voltage V_2 at port 2, and the two currents, I_1 and I_2, can be measured to represent the network response. In this case, V_1 and V_2 are independent variables and I_1 and I_2 are dependent variables, and the network operation can be described by the two equations

$$I_1 = y_{11}V_1 + y_{12}V_2 \tag{C.1}$$

$$I_2 = y_{21}V_1 + y_{22}V_2 \tag{C.2}$$

Here, the four parameters y_{11}, y_{12}, y_{21}, and y_{22} are admittances, and their values completely characterize the linear two-port network.

Depending on which two of the four port variables are used to represent the network excitation, a different set of equations (and a correspondingly different set of parameters) is obtained for characterizing the network. We shall present the four parameter sets commonly used in electronics.

C.1.1 *y* Parameters

The short-circuit admittance (or *y*-parameter) characterization is based on exciting the network by V_1 and V_2, as shown in Fig. C.2(a). The describing equations are Eqs. (C.1) and (C.2). The four admittance parameters can be defined according to their roles in Eqs. (C.1) and (C.2).

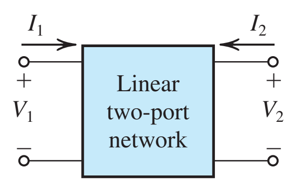


Figure C.1 The reference directions of the four port variables in a linear two-port network.

Specifically, from Eq. (C.1) we see that y_{11} is defined as

$$y_{11} = \left.\frac{I_1}{V_1}\right|_{V_2=0} \tag{C.3}$$

Thus y_{11} is the input admittance at port 1 with port 2 short-circuited. This definition is illustrated in Fig. C.2(b), which also provides a conceptual method for measuring the input short-circuit admittance y_{11}.

The definition of y_{12} can be obtained from Eq. (C.1) as

$$y_{12} = \left.\frac{I_1}{V_2}\right|_{V_1=0} \tag{C.4}$$

Thus y_{12} represents transmission from port 2 to port 1. Since in amplifiers, port 1 represents the input port and port 2 the output port, y_{12} represents internal *feedback* in the network. Figure C.2(c) illustrates the definition of and the method for measuring y_{12}.

The definition of y_{21} can be obtained from Eq. (C.2) as

$$y_{21} = \left.\frac{I_2}{V_1}\right|_{V_2=0} \tag{C.5}$$

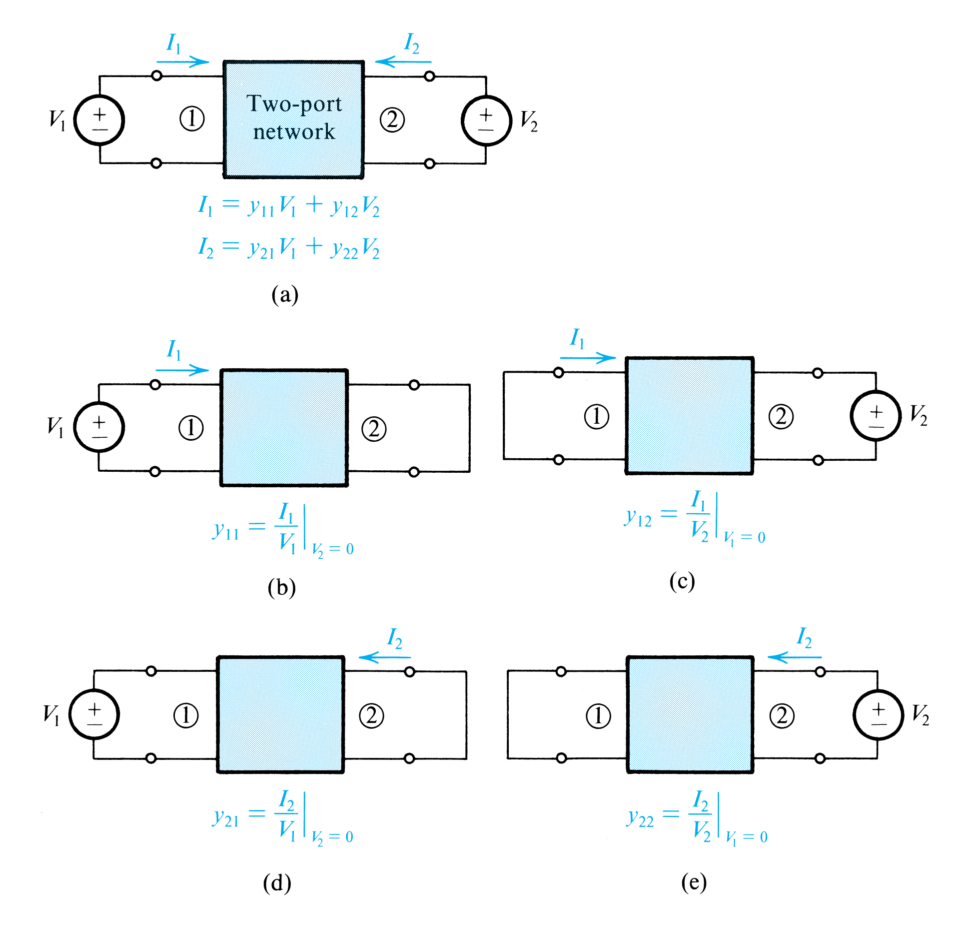


Figure C.2 Definition of and conceptual measurement circuits for the y parameters.

Thus y_{21} represents transmission from port 1 to port 2. If port 1 is the input port and port 2 the output port of an amplifier, then y_{21} provides a measure of the forward gain or transmission. Figure C.2(d) illustrates the definition of and the method for measuring y_{21}.

The parameter y_{22} can be defined, based on Eq. (C.2), as

$$y_{22} = \left.\frac{I_2}{V_2}\right|_{V_1=0} \tag{C.6}$$

Thus y_{22} is the admittance looking into port 2 while port 1 is short-circuited. For amplifiers, y_{22} is the output short-circuit admittance. Figure C.2(e) illustrates the definition of and the method for measuring y_{22}.

C.1.2 *z*

The open-circuit impedance (or z-parameter) characterization of two-port networks is based on exciting the network by I_1 and I_2, as shown in Fig. C.3(a). The describing equations are

$$V_1 = z_{11}I_1 + z_{12}I_2 \tag{C.7}$$

$$V_2 = z_{21}I_1 + z_{22}I_2 \tag{C.8}$$

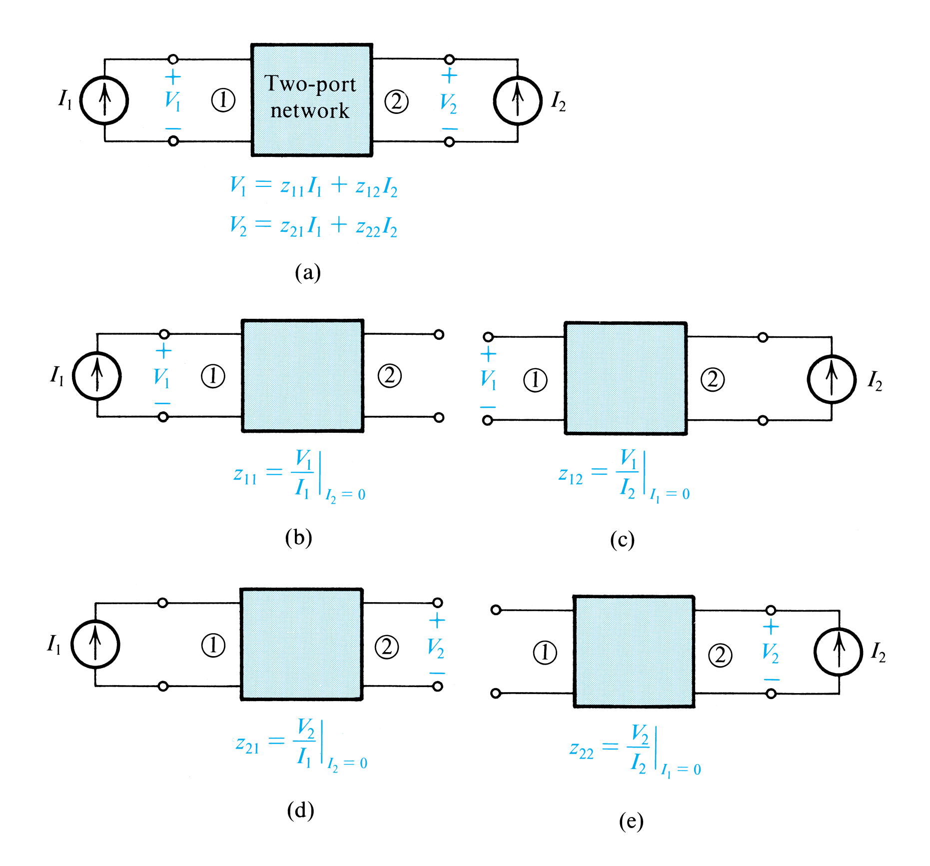


Figure C.3 Definition of and conceptual measurement circuits for the z parameters.

Owing to the duality between the z- and y-parameter characterizations, we shall not give a detailed discussion of z parameters. The definition and the method of measuring each of the four z parameters are given in Fig. C.3.

C.1.3 *h*

The hybrid (or h-parameter) characterization of two-port networks is based on exciting the network by I_1 and V_2, as shown in Fig. C.4(a) (note the reason behind the name *hybrid*). The describing equations are

$$V_1 = h_{11}I_1 + h_{12}V_2 \tag{C.9}$$

$$I_2 = h_{21}I_1 + h_{22}V_2 \tag{C.10}$$

from which the definition of the four h parameters can be obtained as

$$h_{11} = \left.\frac{V_1}{I_1}\right|_{V_2=0} \qquad h_{21} = \left.\frac{I_2}{I_1}\right|_{V_2=0}$$

$$h_{12} = \left.\frac{V_1}{V_2}\right|_{I_0=0} \qquad h_{22} = \left.\frac{I_2}{V_2}\right|_{I_1=0}$$

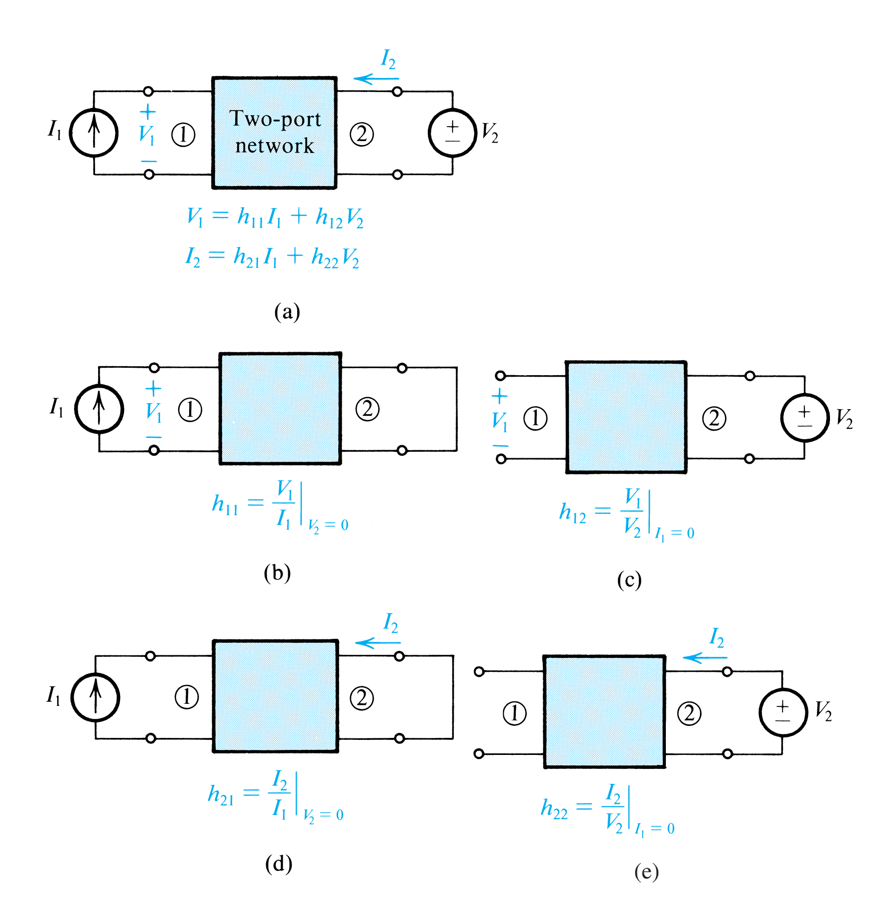


Figure C.4 Definition of and conceptual measurement circuits for the h parameters.

Thus, h_{11} is the input impedance at port 1 with port 2 short-circuited. The parameter h_{12} represents the reverse or feedback voltage ratio of the network, measured with the input port open-circuited. The forward-transmission parameter h_{21} represents the current gain of the network with the output port short-circuited; for this reason, h_{21} is called the *short-circuit current gain*. Finally, h_{22} is the output admittance with the input port open-circuited.

The definitions and conceptual measuring setups of the h parameters are given in Fig. C.4.

C.1.4 *g*

The inverse-hybrid (or g-parameter) characterization of two-port networks is based on excitation of the network by V_1 and I_2, as shown in Fig. C.5(a). The describing equations are

$$I_1 = g_{11}V_1 + g_{12}I_2 \tag{C.11}$$

$$V_2 = g_{21}V_1 + g_{22}I_2 \tag{C.12}$$

The definitions and conceptual measuring setups are given in Fig. C.5.

C.1.5 Equivalent-Circuit Representation

A two-port network can be represented by an equivalent circuit based on the set of parameters used for its characterization. Figure C.6 shows four possible equivalent circuits corresponding

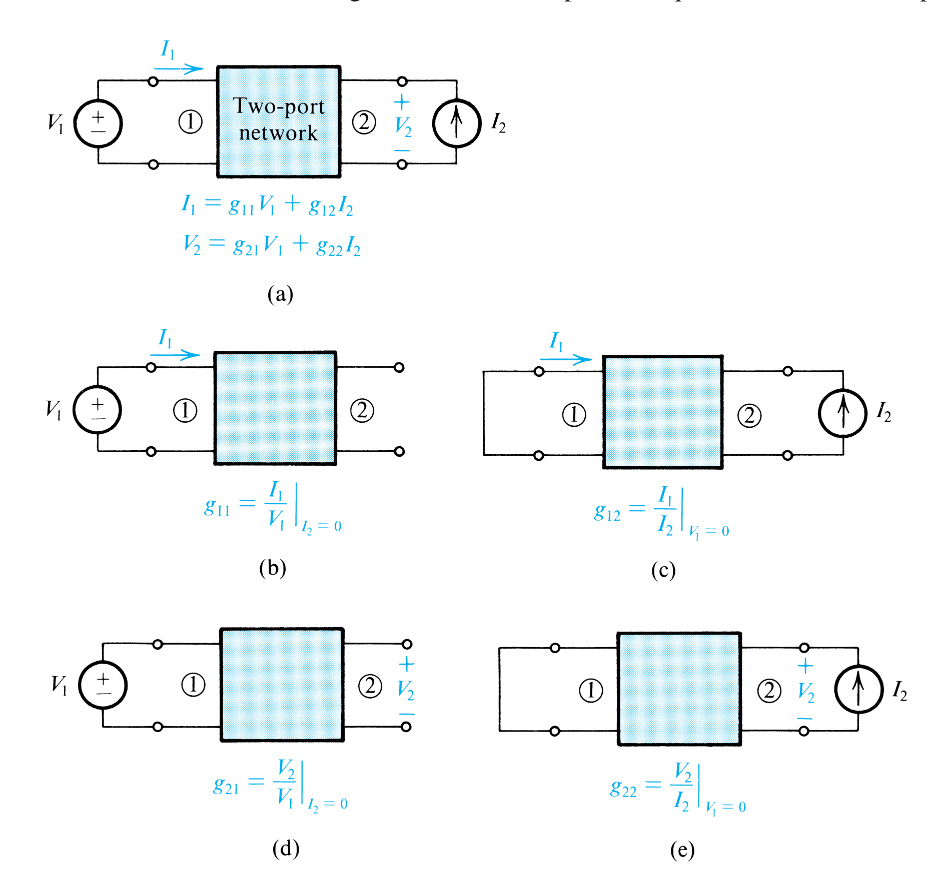


Figure C.5 Definition of and conceptual measurement circuits for the g parameters.

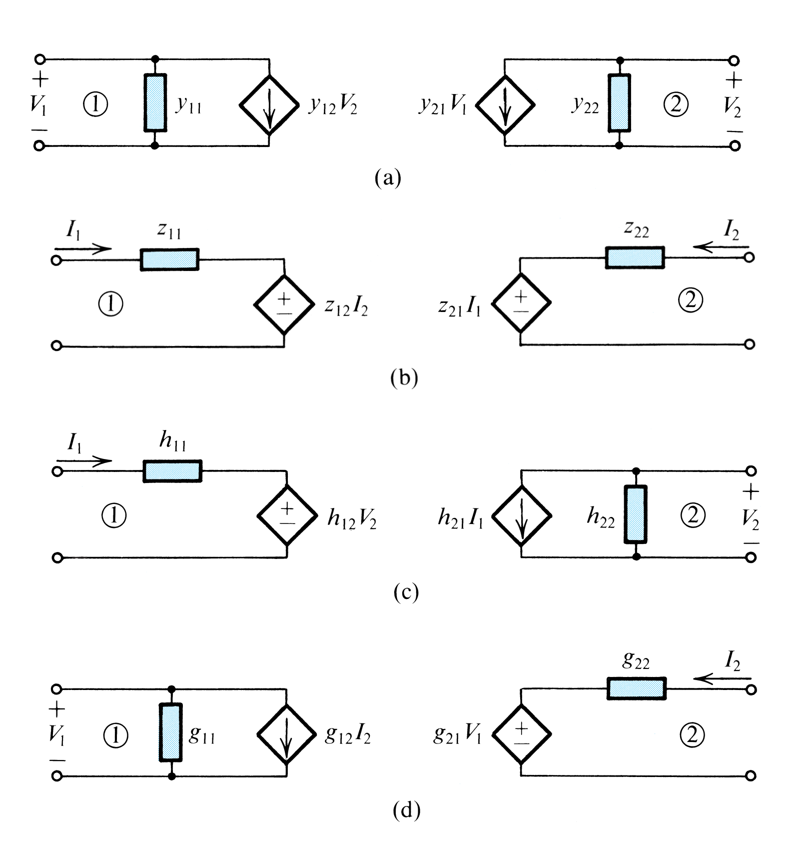


Figure C.6 Equivalent circuits for two-port networks in terms of (**a**) y, (**b**) z, (**c**) h, and (**d**) g parameters.

to the four parameter types just discussed. Each of these equivalent circuits is a direct pictorial representation of the corresponding two equations describing the network in terms of the particular parameter set.

Finally, it should be mentioned that other parameter sets exist for characterizing two-port networks, but these are not discussed or used in this book.

EXERCISE

C.1 Figure EC.1 shows the small-signal, equivalent-circuit model of a transistor. Calculate the values of the h parameters.

Ans. $h_{11} \simeq 2.6\ \text{k}\Omega; h_{12} \simeq 2.5 \times 10^{-4}; h_{21} \simeq 100; h_{22} \simeq 2 \times 10^{-5}\ \Omega$

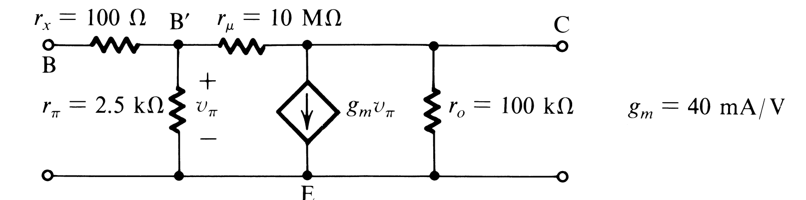


Figure EC.1

C.1 (a) An amplifier characterized by the h-parameter equi-valent circuit of Fig. C.6(c) is fed with a source having a voltage V_s and a resistance R_s, and is loaded in a resistance R_L. Show that its voltage gain is given by

$$\frac{V_2}{V_s} = \frac{-h_{21}}{(h_{11} + R_s)(h_{22} + 1/R_L) - h_{12}h_{21}}$$

(b) Use the expression derived in (a) to find the voltage gain of the transistor in Exercise C.1 for $R_s = 1$ kΩ and $R_L = 10$ kΩ.

C.2 The terminal properties of a two-port network are measured with the following results: With the output short-circuited and an input current of 0.01 mA, the output current is 1.0 mA and the input voltage is 26 mV. With the input open-circuited and a voltage of 10 V applied to the output, the current in the output is 0.2 mA and the voltage measured at the input is 2.5 mV. Find values for the h parameters of this network.

C.3 Figure PC.3 shows the high-frequency equivalent circuit of a BJT. (For simplicity, r_x has been omitted.) Find the y parameters.

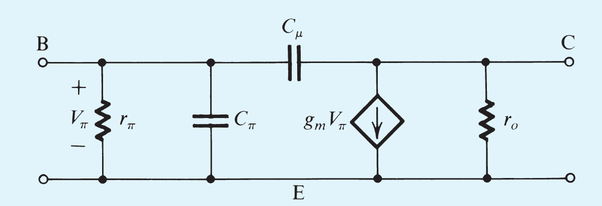


Figure PC.3

APPENDIX D

SOME USEFUL NETWORK THEOREMS

Introduction

In this appendix we review three network theorems that are useful in simplifying the analysis of electronic circuits: Thévenin's theorem, Norton's theorem, and the source-absorption theorem.

D.1 Thévenin's Theorem

Thévenin's theorem is used to represent a part of a network by a voltage source V_t and a series impedance Z_t, as shown in Fig. D.1. Figure D.1(a) shows a network divided into two parts, A and B. In Fig. D.1(b), part A of the network has been replaced by its Thévenin equivalent: a voltage source V_t and a series impedance Z_t. Figure D.1(c) illustrates how V_t is to be determined: Simply open-circuit the two terminals of network A and measure (or calculate) the voltage that appears between these two terminals. To determine Z_t, we reduce all external (i.e., independent) sources in network A to zero by short-circuiting voltage sources and open-circuiting current sources. The impedance Z_t will be equal to the input impedance of network A after this reduction has been performed, as illustrated in Fig. D.1(d).

D.2 Norton's Theorem

Norton's theorem is the *dual* of Thévenin's theorem. It is used to represent a part of a network by a current source I_n and a parallel impedance Z_n, as shown in Fig. D.2. Figure D.2(a) shows a network divided into two parts, A and B. In Fig. D.2(b), part A has been replaced by its Norton's equivalent: a current source I_n and a parallel impedance Z_n. The Norton's current source I_n can be measured (or calculated) as shown in Fig. D.2(c). The terminals of the network being reduced (network A) are shorted, and the current I_n will be equal simply to the short-circuit current. To determine the impedance Z_n, we first reduce the external excitation in network A to zero: That is, we short-circuit independent voltage sources and open-circuit independent current sources. The impedance Z_n will be equal to the input impedance of network A after this source-elimination process has taken place. Thus the Norton impedance Z_n is equal to the Thévenin impedance Z_t. Finally, note that $I_n = V_t/Z$, where $Z = Z_n = Z_t$.

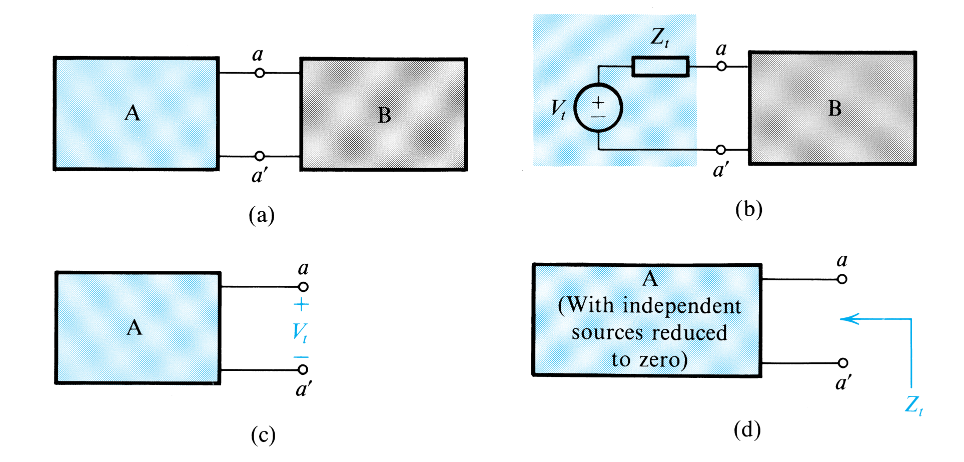


Figure D.1 Thévenin's theorem.

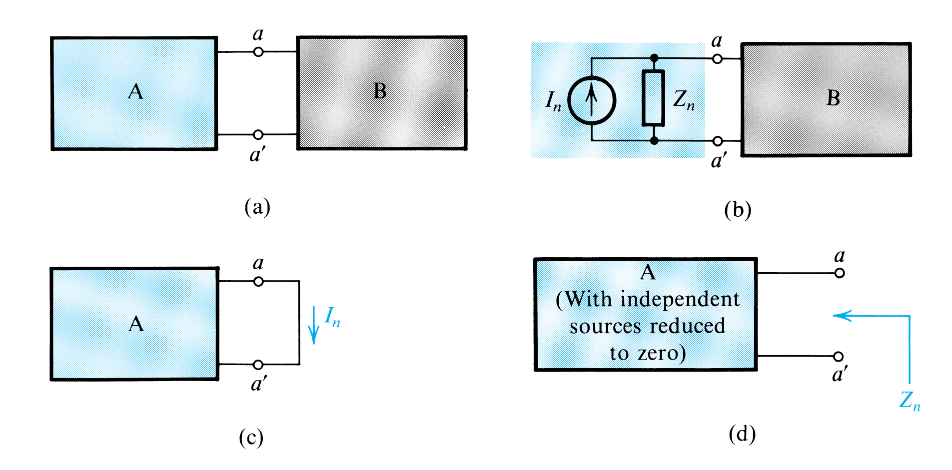


Figure D.2 Norton's theorem.

Example D.1

Figure D.3(a) shows a bipolar junction transistor circuit. The transistor is a three-terminal device with the terminals labeled E (emitter), B (base), and C (collector). As shown, the base is connected to the dc power supply V^+ via the voltage divider composed of R_1 and R_2. The collector is connected to the dc supply V^+ through R_3 and to ground through R_4. To simplify the analysis, we wish to apply Thévenin's theorem to reduce the circuit.

Solution

Thévenin's theorem can be used at the base side to reduce the network composed of V^+, R_1, and R_2 to a dc voltage source V_{BB},

$$V_{BB} = V^+ \frac{R_2}{R_1 + R_2}$$

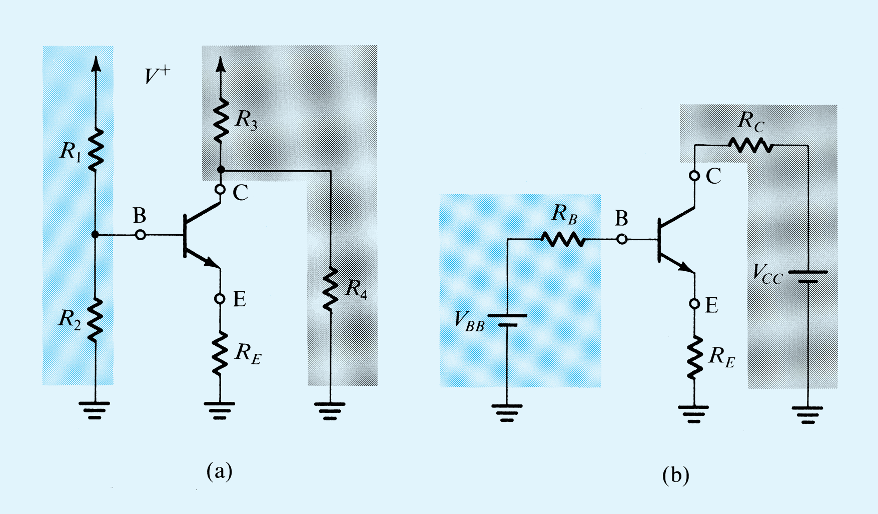


Figure D.3 Thévenin's theorem applied to simplify the circuit of (**a**) to that in (**b**). (See Example D.1.)

and a resistance R_B,

$$R_B = R_1 \parallel R_2$$

where $\parallel$ denotes "in parallel with." At the collector side, Thévenin's theorem can be applied to reduce the network composed of V^+, R_3, and R_4 to a dc voltage source V_{CC},

$$V_{CC} = V^+ \frac{R_4}{R_3 + R_4}$$

and a resistance R_C,

$$R_C = R_3 \parallel R_4$$

The reduced circuit is shown in Fig. D.3(b).

D.3 Source-Absorption Theorem

Consider the situation shown in Fig. D.4. In the course of analyzing a network, we find a controlled current source I_x appearing between two nodes whose voltage difference is the controlling voltage V_x. That is, $I_x = g_m V_x$ where g_m is a conductance. We can replace this controlled source by an impedance $Z_x = V_x/I_x = 1/g_m$, as shown in Fig. D.4, because the current drawn by this impedance will be equal to the current of the controlled source that we have replaced.

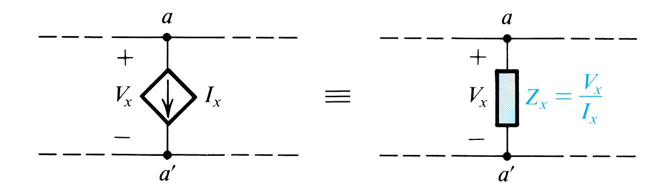


Figure D.4 The source-absorption theorem.

Example D.2

Figure D.5(a) shows the small-signal, equivalent-circuit model of a transistor. We want to find the resistance R_{in} "looking into" the emitter terminal E—that is, the resistance between the emitter and ground—with the base B and collector C grounded.

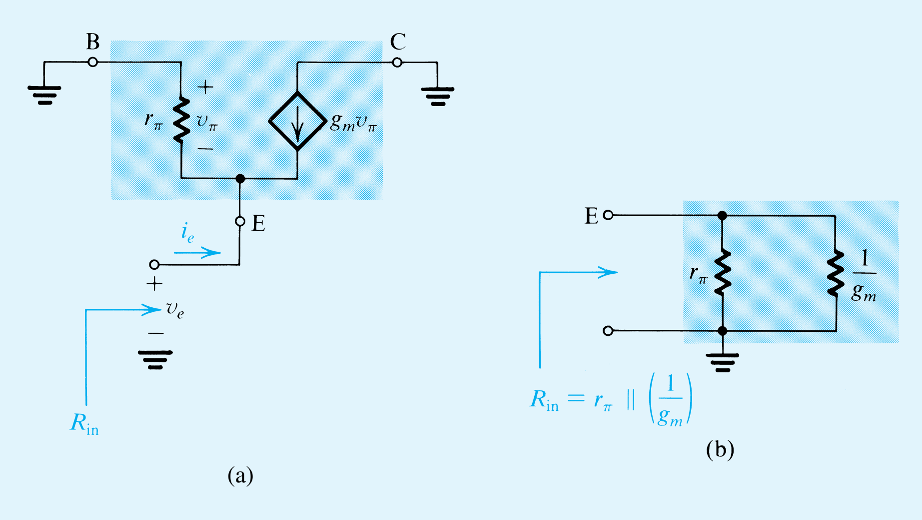


Figure D.5 Circuit for Example D.2.

Solution

From Fig. D.5(a), we see that the voltage v_π will be equal to $-v_e$. Thus, looking between E and ground, we see a resistance r_π in parallel with a current source drawing a current $g_m v_e$ away from terminal E. The latter source can be replaced by a resistance $(1/g_m)$, resulting in the input resistance R_{in} given by

$$R_{in} = r_\pi \parallel (1/g_m)$$

as illustrated in Fig. D.5(b).

EXERCISES

D.1 A source is measured and found to have a 10-V open-circuit voltage and to provide 1 mA into a short circuit. Calculate its Thévenin and Norton equivalent source parameters.
Ans. $V_t = 10$ V; $Z_t = Z_n = 10$ kΩ; $I_n = 1$ mA

D.2 In the circuit shown in Fig. ED.2, the diode has a voltage drop $V_D \simeq 0.7$ V. Use Thévenin's theorem to simplify the circuit and hence calculate the diode current I_D.
Ans. 1 mA

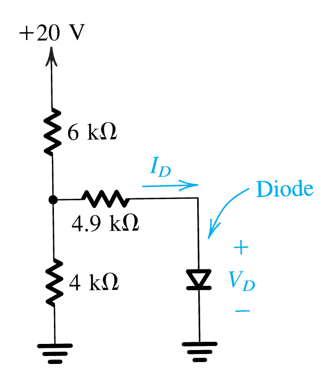


Figure ED.2

D.3 The two-terminal device M in the circuit of Fig. ED.3 has a current $I_M \simeq 1$ mA independent of the voltage V_M across it. Use Norton's theorem to simplify the circuit and hence calculate the voltage V_M.
Ans. 5 V

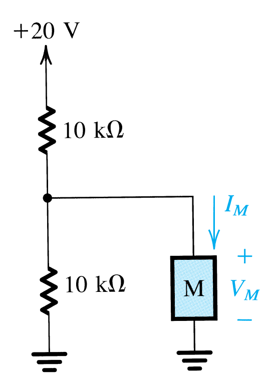


Figure ED.3

PROBLEMS

D.1 Consider the Thévenin equivalent circuit characterized by V_t and Z_t. Find the open-circuit voltage V_{oc} and the short-circuit current I_{sc} (i.e., the current that flows when the terminals are shorted together). Express Z_t in terms of V_{oc} and I_{sc}.

D.2 Repeat Problem D.1 for a Norton equivalent circuit characterized by I_n and Z_n.

D.3 A voltage divider consists of a 9-kΩ resistor connected to +10 V and a resistor of 1 kΩ connected to ground. What is the Thévenin equivalent of this voltage divider? What output voltage results if it is loaded with 1 kΩ? Calculate this two ways: directly and using your Thévenin equivalent.

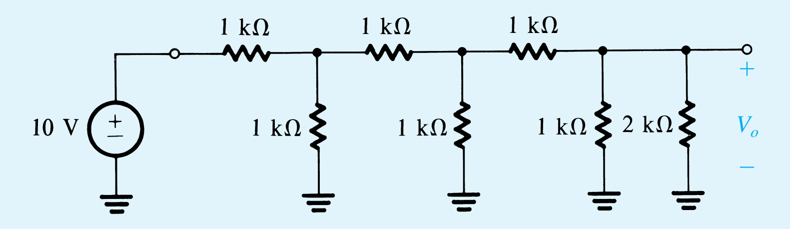


Figure PD.4

D.4 Find the output voltage and output resistance of the circuit shown in Fig. PD.4 by considering a succession of Thévenin equivalent circuits.

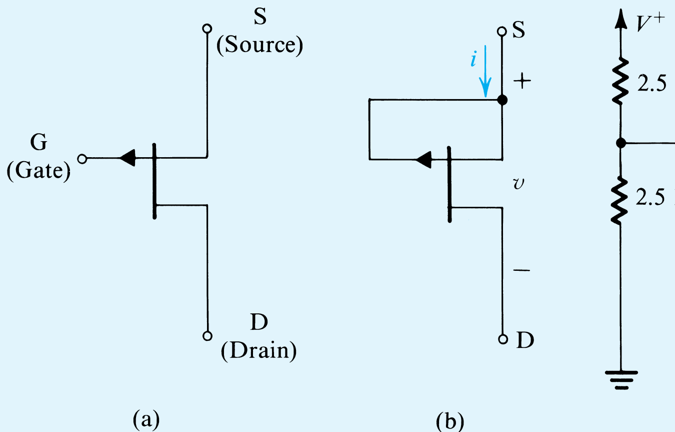
D.5 Repeat Example D.2 with a resistance R_B connected between B and ground in Fig. D.5 (i.e., rather than directly grounding the base B as indicated in Fig. D.5).

D.6 Figure PD.6(a) shows the circuit symbol of a device known as the *p*-channel junction field-effect transistor (JFET). As indicated, the JFET has three terminals. When the gate terminal G is connected to the source terminal S, the two-terminal device shown in Fig. PD.6(b) is obtained. Its i–v characteristic is given by

$$i = I_{DSS}\left[2\frac{v}{V_P} - \left(\frac{v}{V_P}\right)^2\right] \qquad \text{for } v \leq V_P$$

$$i = I_{DSS} \qquad \text{for } v \geq V_P$$

where I_{DSS} and V_P are positive constants for the particular JFET. Now consider the circuit shown in Fig. PD.6(c) and let $V_P = 2$ V and $I_{DSS} = 2$ mA. For $V^+ = 10$ V show that the JFET is operating in the constant-current mode and find the voltage across it. What is the minimum value of V^+ for which this mode of operation is maintained? For $V^+ = 2$ V find the values of I and V.



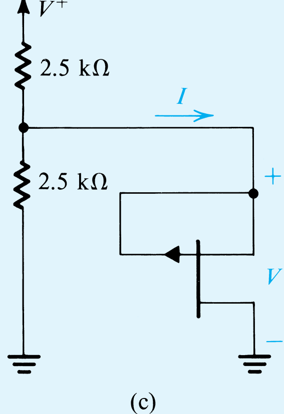


Figure PD.6

APPENDIX E

SINGLE-TIME-CONSTANT CIRCUITS

Introduction

Single-time-constant (STC) circuits are those circuits that are composed of or can be reduced to one reactive component (inductance or capacitance) and one resistance. An STC circuit formed of an inductance L and a resistance R has a time constant $\tau = L/R$. The time constant τ of an STC circuit composed of a capacitance C and a resistance R is given by $\tau = CR$.

Although STC circuits are quite simple, they play an important role in the design and analysis of linear and digital circuits. For instance, the analysis of an amplifier circuit can usually be reduced to the analysis of one or more STC circuits. For this reason, we will review in this appendix the process of evaluating the response of STC circuits to sinusoidal and other input signals such as step and pulse waveforms. The latter signal waveforms are encountered in some amplifier applications but are more important in switching circuits, including digital circuits.

E.1 Evaluating the Time Constant

The first step in the analysis of an STC circuit is to evaluate its time constant τ.

Example E.1

Reduce the circuit in Fig. E.1(a) to an STC circuit, and find its time constant.

Solution

The reduction process is illustrated in Fig. E.1 and consists of repeated applications of Thévenin's theorem. From the final circuit (Fig. E.1c), we obtain the time constant as

$$\tau = C\{R_4 \| [R_3 + (R_1 \| R_2)]\}$$

E.1.1 Rapid Evaluation of τ

In many instances, it will be important to be able to evaluate rapidly the time constant τ of a given STC circuit. A simple method for accomplishing this goal consists first of reducing the excitation to zero; that is, if the excitation is by a voltage source, short it, and if by a current

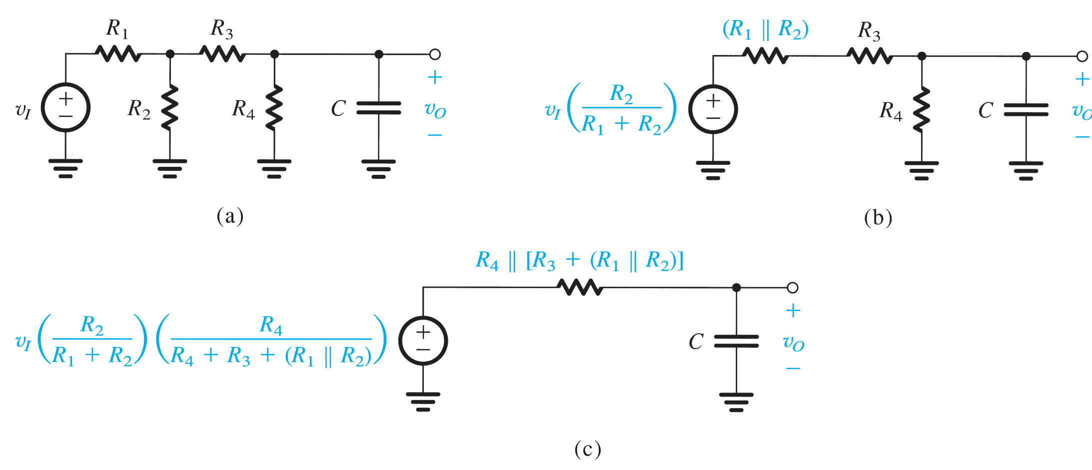


Figure E.1 The reduction of the circuit in **(a)** to the STC circuit in **(c)** through the repeated application of Thévenin's theorem.

source, open it. Then, if the circuit has one reactive component and a number of resistances, "grab hold" of the two terminals of the reactive component (capacitance or inductance) and find the equivalent resistance R_{eq} seen by the component. The time constant is then either L/R_{eq} or CR_{eq}. As an example, in the circuit of Fig. E.1(a), we find that the capacitor C "sees" a resistance R_4 in parallel with the series combination of R_3 and R_2 in parallel with R_1. Thus

$$R_{eq} = R_4 \| [R_3 + (R_2 \| R_1)]$$

and the time constant is CR_{eq}.

In some cases it may be found that the circuit has one resistance and a number of capacitances or inductances. In such a case, the procedure should be inverted; that is, "grab hold" of the resistance terminals and find the equivalent capacitance C_{eq}, or equivalent inductance L_{eq}, seen by this resistance. The time constant is then found as $C_{eq}R$ or L_{eq}/R. This is illustrated in Example E.2.

Example E.2

Find the time constant of the circuit in Fig. E.2.

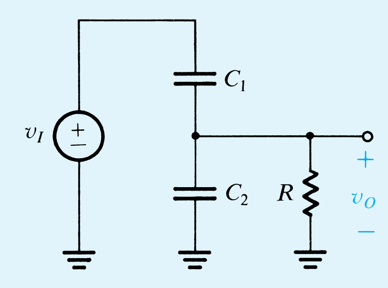


Figure E.2 Circuit for Example E.2.

Solution

After reducing the excitation to zero by short-circuiting the voltage source, we see that the resistance R "sees" an equivalent capacitance $C_1 + C_2$. Thus, the time constant τ is given by

$$\tau = (C_1 + C_2)R$$

Finally, in some cases an STC circuit has more than one resistance and more than one capacitance (or more than one inductance). Such cases require some initial work to simplify the circuit, as illustrated by Example E.3.

Example E.3

Here we show that the response of the circuit in Fig. E.3(a) can be obtained using the method of analysis of STC circuits.

Solution

The analysis steps are illustrated in Fig. E.3. In Fig. E.3(b) we show the circuit excited by two separate but equal voltage sources. The reader should convince himself or herself of the equivalence of the circuits in Fig. E.3(a) and E.3(b). The "trick" employed to obtain the arrangement in Fig. E.3(b) is a very useful one.

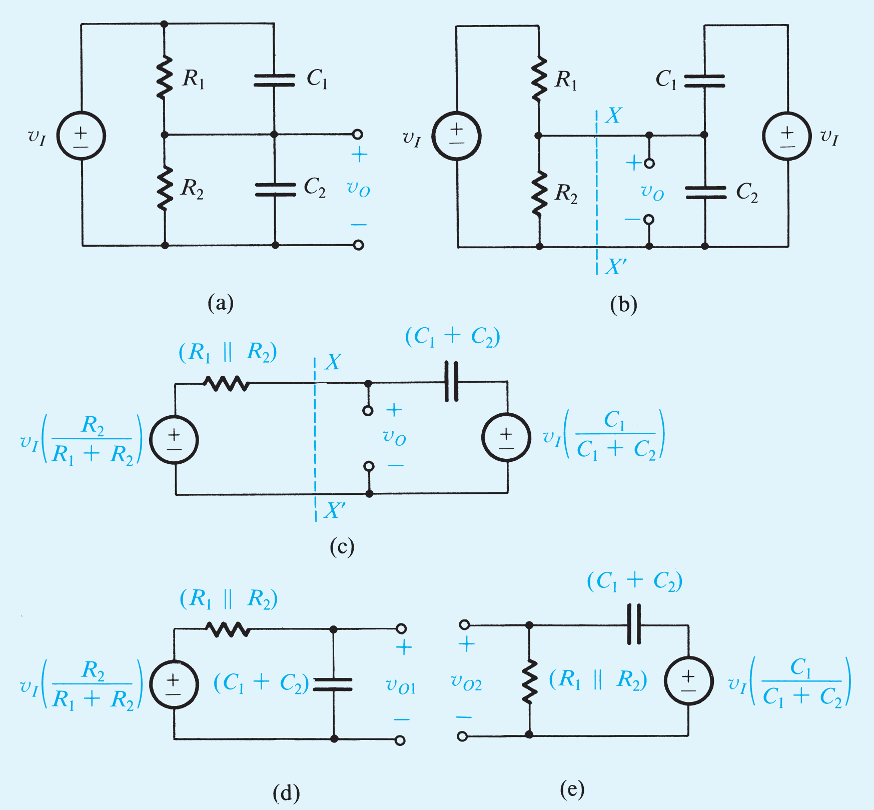


Figure E.3 The response of the circuit in **(a)** can be found by superposition, that is, by summing the responses of the circuits in **(d)** and **(e)**.

Example E.3 *continued*

Application of Thévenin's theorem to the circuit to the left of the line XX' and then to the circuit to the right of that line results in the circuit of Fig. E.3(c). Since this is a linear circuit, the response may be obtained using the principle of superposition. Specifically, the output voltage v_O will be the sum of the two components v_{O1} and v_{O2}. The first component, v_{O1}, is the output due to the left-hand-side voltage source with the other voltage source reduced to zero. The circuit for calculating v_{O1} is shown in Fig. E.3(d). It is an STC circuit with a time constant given by

$$\tau = (C_1 + C_2)(R_1 \parallel R_2)$$

Similarly, the second component v_{O2} is the output obtained with the left-hand-side voltage source reduced to zero. It can be calculated from the circuit of Fig. E.3(e), which is an STC circuit with the same time constant τ.

Finally, it should be observed that the fact that the circuit is an STC one can also be ascertained by setting the independent source v_I in Fig. E.3(a) to zero. Also, the time constant is then immediately obvious.

E.2 Classification of STC Circuits

STC circuits can be classified into two categories, *low-pass* (LP) and *high-pass* (HP) types, with each category displaying distinctly different signal responses. The task of finding whether an STC circuit is of LP or HP type may be accomplished in a number of ways, the simplest of which uses the frequency domain response. Specifically, low-pass circuits pass dc (i.e., signals with zero frequency) and attenuate high frequencies, with the transmission being zero at $\omega = \infty$. Thus, we can test for the circuit type either at $\omega = 0$ or at $\omega = \infty$. At $\omega = 0$ capacitors should be replaced by open circuits ($1/j\omega C = \infty$) and inductors should be replaced by short circuits ($j\omega L = 0$). Then if the output is zero, the circuit is of the high-pass type, while if the output is finite, the circuit is of the low-pass type. Alternatively, we may test at $\omega = \infty$ by replacing capacitors with short circuits ($1/j\omega C = 0$) and inductors with open circuits ($j\omega L = \infty$). Then if the output is finite, the circuit is of the HP type, whereas if the output is zero, the circuit is of the LP type. In Table E.1, which provides a summary of these results, s.c. stands for short circuit and o.c. for open circuit.

Figure E.4 shows examples of low-pass STC circuits, and Fig. E.5 shows examples of high-pass STC circuits. For each circuit we have indicated the input and output variables of interest. Note that a given circuit can be of either category, depending on the input and output variables. The reader is urged to verify, using the rules of Table E.1, that the circuits of Figs. E.4 and E.5 are correctly classified.

Table E.1 Rules for finding the type of STC Circuit

Test at	Replace	Circuit is LP if	Circuit is HP if
$\omega = 0$	C by o.c. L by s.c.	output is finite	output is zero
$\omega = \infty$	C by s.c. L by o.c.	output is zero	output is finite

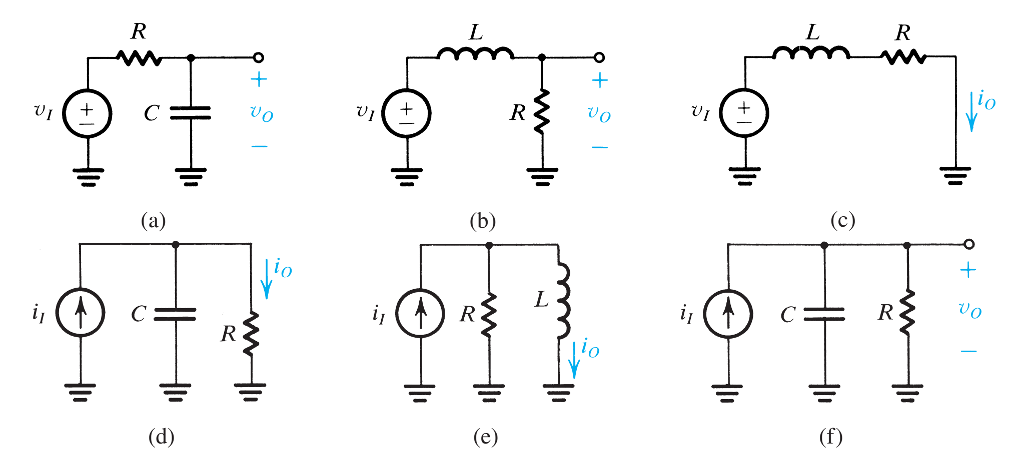


Figure E.4 STC circuits of the low-pass type.

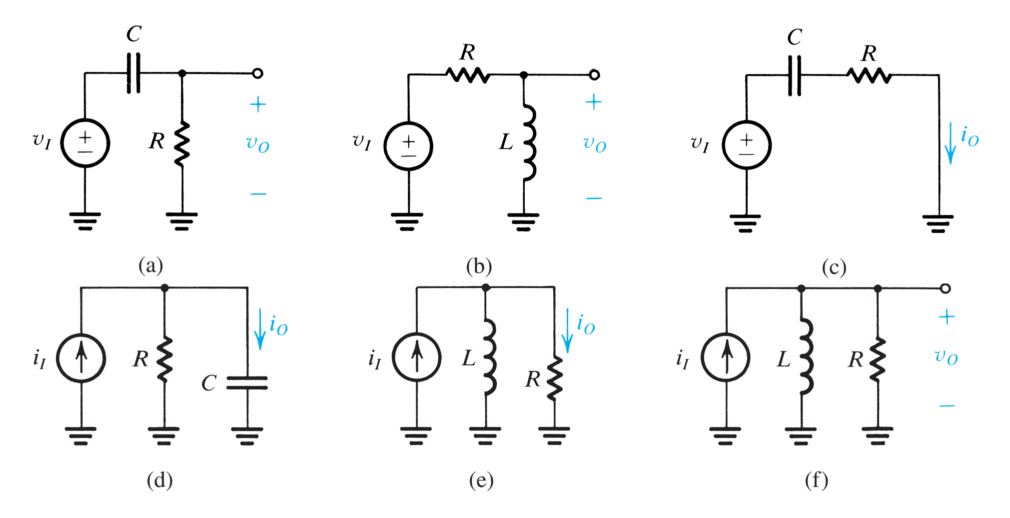


Figure E.5 STC circuits of the high-pass type.

EXERCISES

E.1 Find the time constants for the circuits shown in Fig. EE.1.

Ans. (a) $\frac{(L_1 \| L_2)}{R}$; (b) $\frac{(L_1 \| L_2)}{(R_1 \| R_2)}$

(a) (b)

Figure EE.1

E.2 Classify the following circuits as STC high-pass or low-pass: Fig. E.4(a) with output i_o in C to ground; Fig. E.4(b) with output i_o in R to ground; Fig. E.4(d) with output i_o in C to ground; Fig. E.4(e) with output i_o in R to ground; Fig. E.5(b) with output i_o in L to ground; and Fig. E.5(d) with output v_o across C.

Ans. HP; LP; HP; HP; LP; LP

E.3 Frequency Response of STC Circuits

E.3.1 Low-Pass Circuits

The transfer function $T(s)$ of an STC low-pass circuit can always be written in the form

$$T(s) = \frac{K}{1 + (s/\omega_0)} \tag{E.1}$$

which, for physical frequencies, where $s = j\omega$, becomes

$$T(j\omega) = \frac{K}{1 + j(\omega/\omega_0)} \tag{E.2}$$

where K is the magnitude of the transfer function at $\omega = 0$ (dc) and ω_0 is defined by

$$\omega_0 = 1/\tau$$

with τ being the time constant. Thus the magnitude response is given by

$$|T(j\omega)| = \frac{K}{\sqrt{1 + (\omega/\omega_0)^2}} \tag{E.3}$$

and the phase response is given by

$$\phi(\omega) = -\tan^{-1}(\omega/\omega_0) \tag{E.4}$$

Figure E.6 sketches the magnitude and phase responses for an STC low-pass circuit. The magnitude response shown in Fig. E.6(a) is simply a graph of the function in Eq. (E.3). The magnitude is normalized with respect to the dc gain K and is expressed in decibels; that is, the plot is for $20 \log|T(j\omega)/K|$, with a logarithmic scale used for the frequency axis. Furthermore, the frequency variable has been normalized with respect to ω_0. As shown, the magnitude curve is closely defined by two straight-line asymptotes. The low-frequency asymptote is a horizontal straight line at 0 dB. To find the slope of the high-frequency asymptote, consider Eq. (E.3) and let $\omega/\omega_0 \gg 1$, resulting in

$$|T(j\omega)| \simeq K\frac{\omega_0}{\omega}$$

It follows that if ω doubles in value, the magnitude is halved. On a logarithmic frequency axis, doublings of ω represent equally spaced points, with each interval called an *octave*. Halving the magnitude function corresponds to a 6-dB reduction in transmission ($20 \log 0.5 = -6$ dB). Thus the slope of the high-frequency asymptote is –6 dB/octave. This can be equivalently expressed as –20 dB/decade, where “decade” indicates an increase in frequency by a factor of 10.

The two straight-line asymptotes of the magnitude–response curve meet at the “corner frequency” or “break frequency” ω_0. The difference between the actual magnitude–response

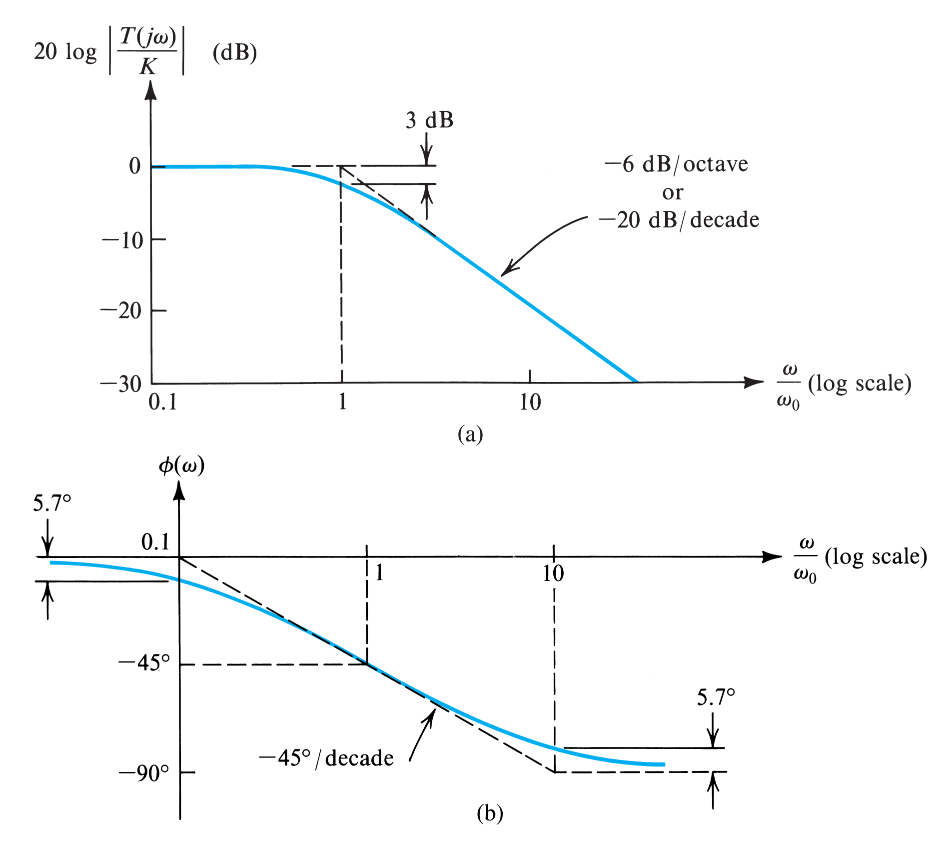


Figure E.6 **(a)** Magnitude and **(b)** phase response of STC circuits of the low-pass type.

curve and the asymptotic response is largest at the corner frequency, where its value is 3 dB. To verify that this value is correct, simply substitute $\omega = \omega_0$ in Eq. (E.3) to obtain

$$|T(j\omega_0)| = K/\sqrt{2}$$

Thus at $\omega = \omega_0$, the gain drops by a factor of $\sqrt{2}$ relative to the dc gain, which corresponds to a 3-dB reduction in gain. The corner frequency ω_0 is appropriately referred to as the 3-dB frequency.

Similar to the magnitude response, the phase–response curve, shown in Fig. E.6(b), is closely defined by straight-line asymptotes. Note that at the corner frequency the phase is –45°, and that for $\omega \gg \omega_0$ the phase approaches –90°. Also note that the –45°/decade straight line approximates the phase function, with a maximum error of 5.7°, over the frequency range $0.1\omega_0$ to $10\omega_0$.

Example E.4

Consider the circuit shown in Fig. E.7(a), where an ideal voltage amplifier of gain $\mu = -100$ has a small (10-pF) capacitance connected in its feedback path. The amplifier is fed by a voltage source having a source resistance of 100 kΩ. Show that the frequency response V_o/V_s of this amplifier is equivalent to that of an STC circuit, and sketch the magnitude response.

Example E.4 *continued*

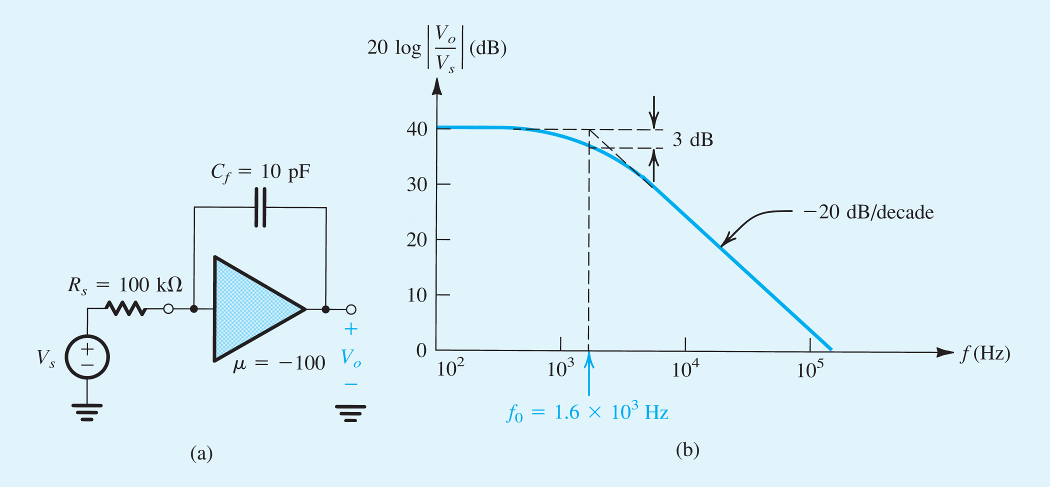


Figure E.7 **(a)** An amplifier circuit and **(b)** a sketch of the magnitude of its transfer function.

Solution

Direct analysis of the circuit in Fig. E.7(a) results in the transfer function

$$\frac{V_o}{V_s} = \frac{\mu}{1 + sRC_f(-\mu + 1)}$$

which can be seen to be that of a low-pass STC circuit with a dc gain $\mu = -100$ (or, equivalently, 40 dB) and a time constant $\tau = RC_f(-\mu + 1) = 100 \times 10^3 \times 10 \times 10^{-12} \times 101 \simeq 10^{-4}$ s, which corresponds to a frequency $\omega_0 = 1/\tau = 10^4$ rad/s. The magnitude response is sketched in Fig. E.7(b).

E.3.2 High-Pass Circuits

The transfer function $T(s)$ of an STC high-pass circuit can always be expressed in the form

$$T(s) = \frac{Ks}{s + \omega_0} \tag{E.5}$$

which for physical frequencies $s = j\omega$ becomes

$$T(j\omega) = \frac{K}{1 - j\omega_0/\omega} \tag{E.6}$$

where K denotes the gain as s or ω approaches infinity and ω_0 is the inverse of the time constant τ,

$$\omega_0 = 1/\tau$$

The magnitude response

$$|T(j\omega)| = \frac{K}{\sqrt{1 + (\omega_0/\omega)^2}} \tag{E.7}$$

and the phase response

$$\phi(\omega) = \tan^{-1}(\omega_0/\omega) \tag{E.8}$$

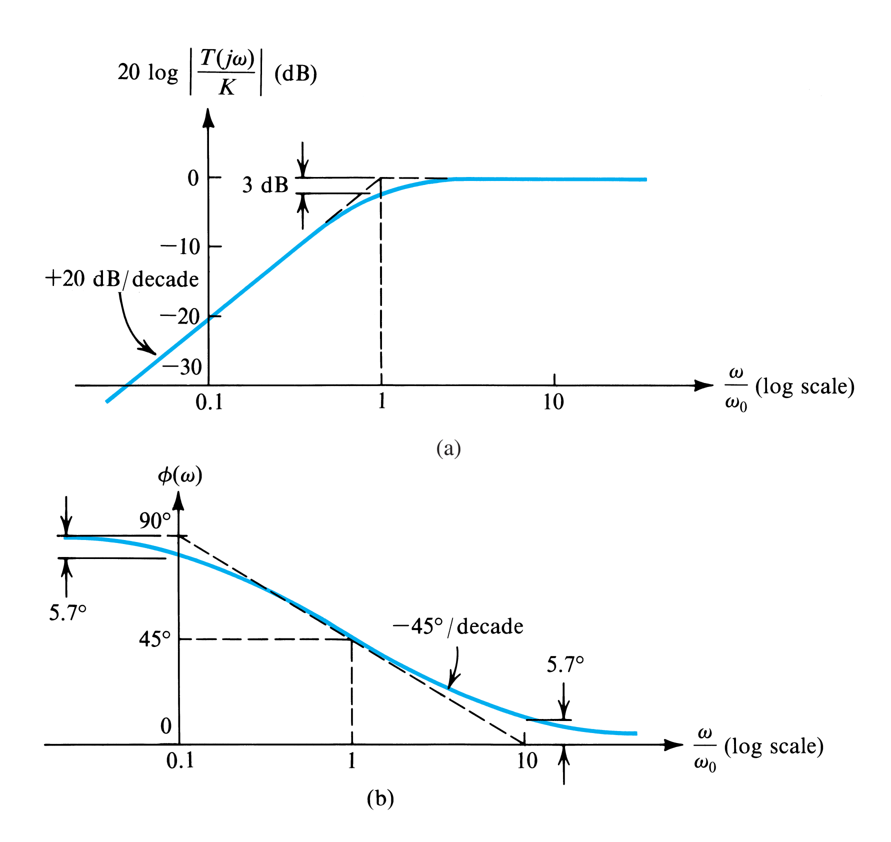


Figure E.8 **(a)** Magnitude and **(b)** phase response of STC circuits of the high-pass type.

are sketched in Fig. E.8. As in the low-pass case, the magnitude and phase curves are well defined by straight-line asymptotes. Because of the similarity (or, more appropriately, duality) with the low-pass case, no further explanation will be given.

EXERCISES

E.3 Find the dc transmission, the corner frequency f_0, and the transmission at $f = 2$ MHz for the low-pass STC circuit shown in Fig. EE.3.

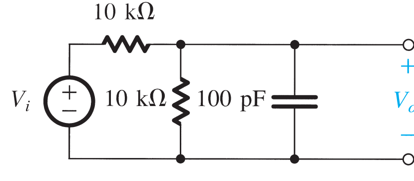


Figure EE.3

Ans. –6 dB; 318 kHz; –22 dB

E.4 Find the transfer function $T(s)$ of the circuit in Fig. E.2. What type of STC network is it?

Ans. $T(s) = \dfrac{C_1}{C_1 + C_2} \dfrac{s}{s + [1/(C_1 + C_2)R]}$; HP

E.5 For the situation discussed in Exercise E.4, if $R = 10\ \text{k}\Omega$, find the capacitor values that result in the circuit having a high-frequency transmission of 0.5 V/V and a corner frequency $\omega_0 = 10$ rad/s.
Ans. $C_1 = C_2 = 5\ \mu\text{F}$

E.6 Find the high-frequency gain, the 3-dB frequency f_0, and the gain at $f = 1$ Hz of the capacitively coupled amplifier shown in Fig. EE.6. Assume the voltage amplifier to be ideal.

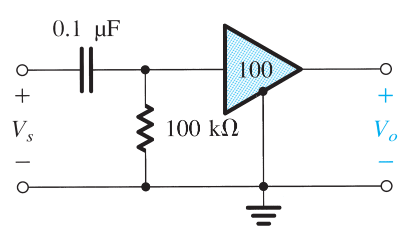


Figure EE.6

Ans. 40 dB; 15.9 Hz; 16 dB

E.4 Step Response of STC Circuits

In this section we consider the response of STC circuits to the step-function signal shown in Fig. E.9. Knowledge of the step response enables rapid evaluation of the response to other switching-signal waveforms, such as pulses and square waves.

E.4.1 Low-Pass Circuits

In response to an input step signal of height S, a low-pass STC circuit (with a dc gain $K = 1$) produces the waveform shown in Fig. E.10. Note that while the input rises from 0 to S at $t = 0$, the output does not respond immediately to this transient and simply begins to rise exponentially toward the *final* dc value of the input, S. In the long term—that is, for $t \gg \tau$ — the output approaches the dc value S, a manifestation of the fact that low-pass circuits faithfully pass dc.

The equation of the output waveform can be obtained from the expression

$$y(t) = Y_\infty - (Y_\infty - Y_{0+})e^{-t/\tau} \tag{E.9}$$

where Y_∞ denotes the *final* value or the value toward which the output is heading and Y_{0+} denotes the value of the output immediately after $t = 0$. This equation states that *the output at*

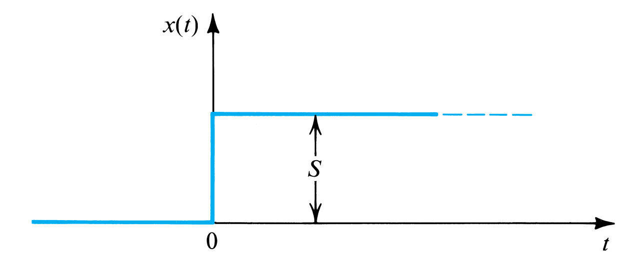


Figure E.9 A step-function signal of height S.

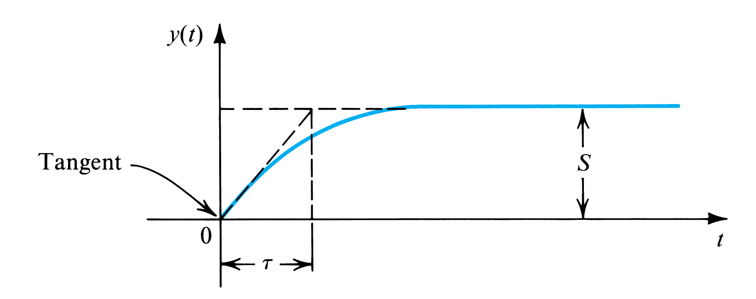


Figure E.10 The output $y(t)$ of a low-pass STC circuit excited by a step of height S.

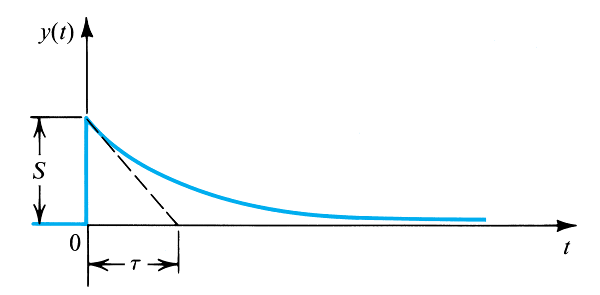


Figure E.11 The output $y(t)$ of a high-pass STC circuit excited by a step of height S.

any time t is equal to the difference between the final value Y_∞ and a gap that has an initial value of $Y_\infty - Y_{0+}$ and is "shrinking" exponentially. In our case, $Y_\infty = S$ and $Y_{0+} = 0$; thus,

$$y(t) = S\left(1 - e^{-t/\tau}\right) \tag{E.10}$$

The reader's attention is drawn to the slope of the tangent to $y(t)$ at $t = 0$, which is indicated in Fig. E.10.

E.4.2 High-Pass Circuits

The response of an STC high-pass circuit (with a high-frequency gain $K = 1$) to an input step of height S is shown in Fig. E.11. The high-pass circuit faithfully transmits the transient part of the input signal (the step change) but blocks the dc. Thus the output at $t = 0$ follows the input,

$$Y_{0+} = S$$

and then it decays toward zero,

$$Y_\infty = 0$$

Substituting for Y_{0+} and Y_∞ in Eq. (E.9) results in the output $y(t)$,

$$y(t) = Se^{-t/\tau} \tag{E.11}$$

The reader's attention is drawn to the slope of the tangent to $y(t)$ at $t = 0$, indicated in Fig. E.11.

Example E.5

This example is a continuation of the problem considered in Example E.3. For an input v_I that is a 10-V step, find the condition under which the output v_O is a perfect step.

Solution

Following the analysis in Example E.3, which is illustrated in Fig. E.3, we have

$$v_{O1} = k_r\left[10\left(1 - e^{-t/\tau}\right)\right]$$

where

$$k_r \equiv \frac{R_2}{R_1 + R_2}$$

and

$$v_{O2} = k_c(10e^{-t/\tau})$$

where

$$k_c \equiv \frac{C_1}{C_1 + C_2}$$

and

$$\tau = \left(C_1 + C_2\right)\left(R_1 \parallel R_2\right)$$

Thus

$$\begin{aligned} v_O &= v_{O1} + v_{O2} \\ &= 10k_r + 10e^{-t/\tau}\left(k_c - k_r\right) \end{aligned}$$

It follows that the output can be made a perfect step of height $10k_r$ volts if we arrange that

$$k_c = k_r$$

that is, if the resistive voltage divider ratio is made equal to the capacitive voltage divider ratio.

This example illustrates an important technique, namely, that of the "compensated attenuator." An application of this technique is found in the design of the oscilloscope probe. The oscilloscope probe problem is investigated in Problem E.3.

EXERCISES

E.7 For the circuit of Fig. E.4(f), find v_O if i_I is a 3-mA step, $R = 1$ kΩ, and $C = 100$ pF.
Ans. $3(1 - e^{-10^7 t})$

E.8 In the circuit of Fig. E.5(f), find $v_O(t)$ if i_I is a 2-mA step, $R = 2$ kΩ, and $L = 10$ µH.
Ans. $4e^{-2\times 10^8 t}$

E.9 The amplifier circuit of Fig. EE.6 is fed with a signal source that delivers a 20-mV step. If the source resistance is 100 kΩ, find the time constant τ and $v_O(t)$.
Ans. $\tau = 2 \times 10^{-2}$ s; $v_O(t) = 1 \times e^{-50t}$

E.10 For the circuit in Fig. E.2 with $C_1 = C_2 = 0.5$ µF, $R = 1$ MΩ, find $v_O(t)$ if $v_I(t)$ is a 10-V step.
Ans. $5e^{-t}$

E.11 Show that the area under the exponential of Fig. E.11 is equal to that of the rectangle of height S and width τ.

E.5 Pulse Response of STC Circuits

Figure E.12 shows a pulse signal whose height is P and whose width is T. We wish to find the response of STC circuits to input signals of this form. Note at the outset that a pulse can be considered as the sum of two steps: a positive one of height P occurring at $t = 0$ and a negative one of height P occurring at $t = T$. Thus, the response of a linear circuit to the pulse signal can be obtained by summing the responses to the two step signals.

E.5.1 Low-Pass Circuits

Figure E.13(a) shows the response of a low-pass STC circuit (having unity dc gain) to an input pulse of the form shown in Fig. E.12. In this case, we have assumed that the time constant τ is in the same range as the pulse width T. As shown, the LP circuit does not respond immediately to the step change at the leading edge of the pulse; rather, the output starts to rise exponentially toward a final value of P. This exponential rise, however, will be stopped at time $t = T$, that is, at the trailing edge of the pulse when the input undergoes a negative step change. Again, the output will respond by starting an exponential decay toward the final value of the input, which is zero. Finally, note that the area under the output waveform will be equal to the area under the input pulse waveform, since the LP circuit faithfully passes dc.

A low-pass effect usually occurs when a pulse signal from one part of an electronic system is connected to another. The low-pass circuit in this case is formed by the output resistance (Thévenin's equivalent resistance) of the system part from which the signal originates and the input capacitance of the system part to which the signal is fed. This unavoidable low-pass filter will cause distortion—of the type shown in Fig. E.13(a)—of the pulse signal. In a well-designed system such distortion is kept to a low value by arranging that the time constant τ be much smaller than the pulse width T. In this case, the result will be a slight rounding of the pulse edges, as shown in Fig. E.13(b). Note, however, that the edges are still exponential.

The distortion of a pulse signal by a parasitic (i.e., unwanted) low-pass circuit is measured by its *rise time* and *fall time*. The rise time is conventionally defined as the time taken by the amplitude to increase from 10% to 90% of the final value. Similarly, the fall time is the time during which the pulse amplitude falls from 90% to 10% of the maximum value. These definitions are illustrated in Fig. E.13(b). By use of the exponential equations of the rising and falling edges of the output waveform, it can be easily shown that

$$t_r = t_f \simeq 2.2\tau \tag{E.12}$$

which can be also expressed in terms of $f_0 = \omega_0/2\pi = 1/2\pi\tau$ as

$$t_r = t_f \simeq \frac{0.35}{f_0} \tag{E.13}$$

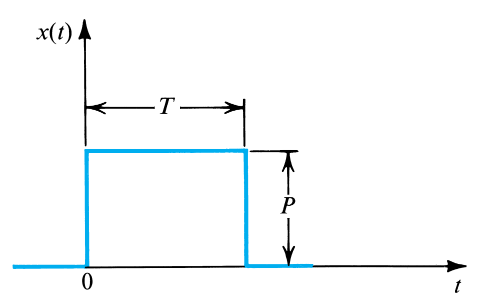


Figure E.12 A pulse signal with height P and width T.

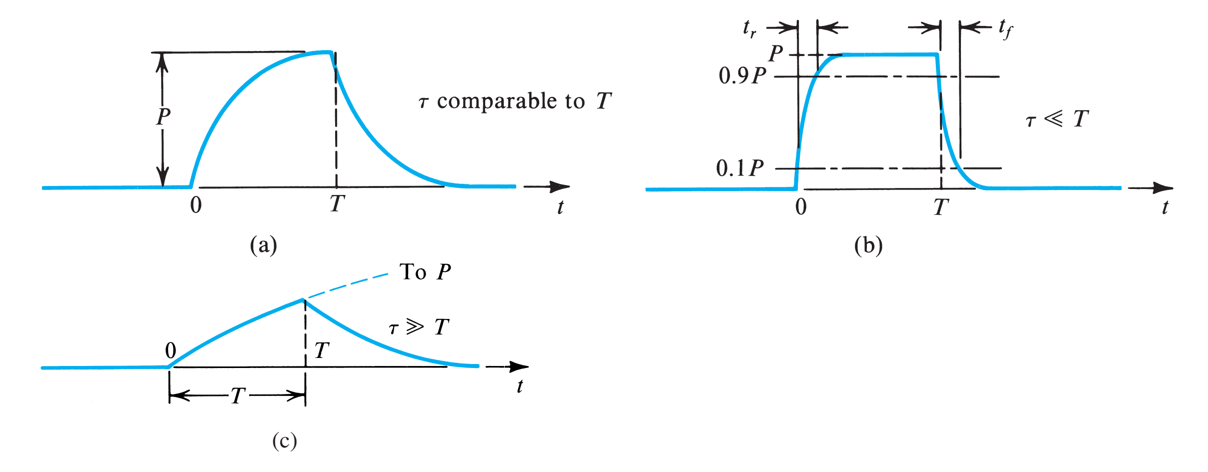


Figure E.13 Pulse responses of three STC low-pass circuits.

Finally, we note that the effect of the parasitic low-pass circuits that are always present in a system is to "slow down" the operation of the system: To keep the signal distortion within acceptable limits, one has to use a relatively long pulse width (for a given low-pass time constant).

The other extreme case—namely, when τ is much larger than T, is illustrated in Fig. E.13(c). As shown, the output waveform rises exponentially toward the level P. However, since $\tau \gg T$, the value reached at $t = T$ will be much smaller than P. At $t = T$, the output waveform starts its exponential decay toward zero. Note that in this case the output waveform bears little resemblance to the input pulse. Also note that because $\tau \gg T$, the portion of the exponential curve from $t = 0$ to $t = T$ is almost linear. Since the slope of this linear curve is proportional to the height of the input pulse, we see that the output waveform approximates the time integral of the input pulse. That is, a low-pass network with a large time constant approximates the operation of an *integrator*.

E.5.2 High-Pass Circuits

Figure E.14(a) shows the output of an STC HP circuit (with unity high-frequency gain) excited by the input pulse of Fig. E.12, assuming that τ and T are comparable in value. As shown, the step transition at the leading edge of the input pulse is faithfully reproduced at the output of the HP circuit. However, since the HP circuit blocks dc, the output waveform immediately starts an exponential decay toward zero. This decay process is stopped at $t = T$, when the negative step transition of the input occurs and the HP circuit faithfully reproduces it. Thus, at $t = T$ the output waveform exhibits an *undershoot*. Then it starts an exponential decay toward zero. Finally, note that the area of the output waveform above the zero axis will be equal to that below the axis for a total average area of zero, consistent with the fact that HP circuits block dc.

In many applications, an STC high-pass circuit is used to couple a pulse from one part of a system to another part. In such an application, it is necessary to keep the distortion in the pulse shape as small as possible. This can be accomplished by selecting the time constant τ to be much longer than the pulse width T. If this is indeed the case, the loss in amplitude during the pulse period T will be very small, as shown in Fig. E.14(b). Nevertheless, the output waveform still swings negatively, and the area under the negative portion will be equal to that under the positive portion.

Consider the waveform in Fig. E.14(b). Since τ is much larger than T, it follows that the portion of the exponential curve from $t = 0$ to $t = T$ will be almost linear and that its slope

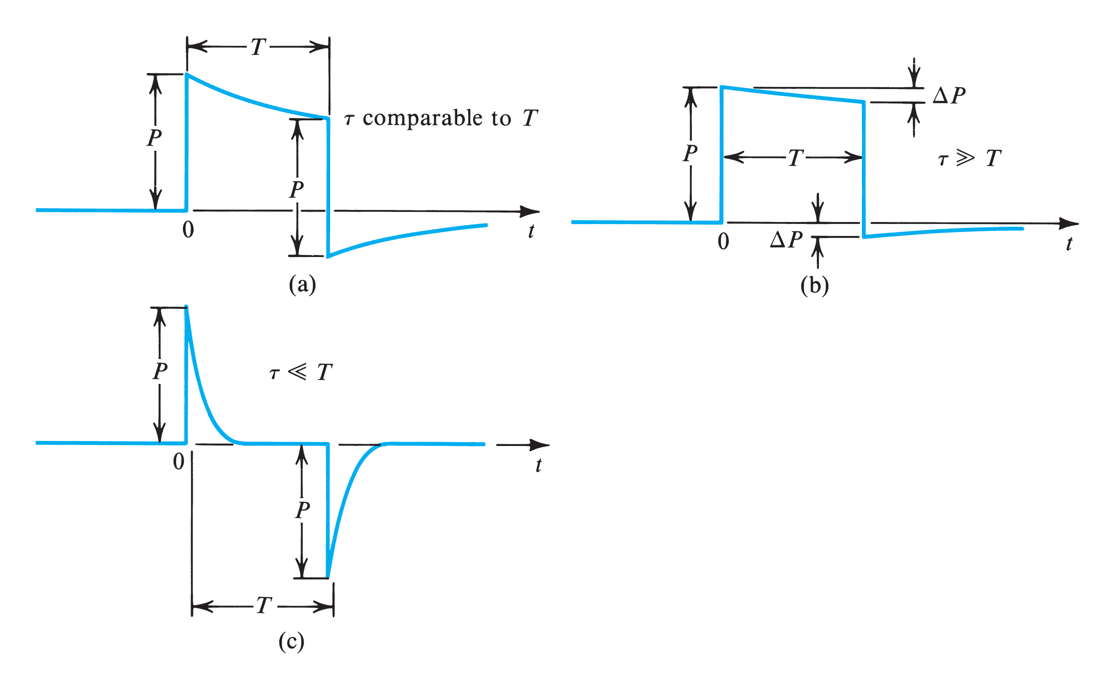


Figure E.14 Pulse responses of three STC high-pass circuits.

will be equal to the slope of the exponential curve at $t = 0$, which is P/τ. We can use this value of the slope to determine the loss in amplitude ΔP as

$$\Delta P \simeq \frac{P}{\tau} T \tag{E.14}$$

The distortion effect of the high-pass circuit on the input pulse is usually specified in terms of the per-unit or percentage loss in pulse height. This quantity is taken as an indication of the "sag" in the output pulse,

$$\text{Percentage sag} \equiv \frac{\Delta P}{P} \times 100 \tag{E.15}$$

Thus

$$\text{Percentage sag} = \frac{T}{\tau} \times 100 \tag{E.16}$$

Finally, note that the magnitude of the undershoot at $t = T$ is equal to ΔP.

The other extreme case—namely, $\tau \ll T$—is illustrated in Fig. E.14(c). In this case, the exponential decay is quite rapid, resulting in the output becoming almost zero shortly beyond the leading edge of the pulse. At the trailing edge of the pulse, the output swings negatively by an amount almost equal to the pulse height P. Then the waveform decays rapidly to zero. As seen from Fig. E.14(c), the output waveform bears no resemblance to the input pulse. It consists of two spikes: a positive one at the leading edge and a negative one at the trailing edge. Note that the output waveform is approximately equal to the time derivative of the input pulse. That is, for $\tau \ll T$, an STC high-pass circuit approximates a *differentiator*. However, the resulting differentiator is not an ideal one; an ideal differentiator would produce two impulses. Nevertheless, high-pass STC circuits with short time constants are employed in some applications to produce sharp pulses or spikes at the transitions of an input waveform.

EXERCISES

E.12 Find the rise and fall times of a 1-μs pulse after it has passed through a low-pass RC circuit with a corner frequency of 10 MHz.
Ans. 35 ns

E.13 Consider the pulse response of a low-pass STC circuit, as shown in Fig. E.13(c). If $\tau = 100T$, find the output voltage at $t = T$. Also, find the difference in the slope of the rising portion of the output waveform at $t = 0$ and $t = T$ (expressed as a percentage of the slope at $t = 0$).
Ans. $0.01P$; 1%

E.14 The output of an amplifier stage is connected to the input of another stage via a capacitance C. If the first stage has an output resistance of 10 kΩ, and the second stage has an input resistance of 40 kΩ, find the minimum value of C such that a 10-μs pulse exhibits less than 1% sag.
Ans. 0.02 μF

E.15 A high-pass STC circuit with a time constant of 100 μs is excited by a pulse of 1-V height and 100-μs width. Calculate the value of the undershoot in the output waveform.
Ans. 0.632 V

PROBLEMS

E.1 Consider the circuit of Fig. E.3(a) and the equivalent shown in (d) and (e). There, the output, $v_O = v_{O1} + v_{O2}$, is the sum of outputs of a low-pass and a high-pass circuit, each with the time constant $\tau = (C_1 + C_2)(R_1 \| R_2)$. What is the condition that makes the contribution of the low-pass circuit at zero frequency equal to the contribution of the high-pass circuit at infinite frequency? Show that this condition can be expressed as $C_1R_1 = C_2R_2$. If this condition applies, sketch $|V_o/V_i|$ versus frequency for the case $R_1 = R_2$.

E.2 Use the voltage divider rule to find the transfer function $V_o(s)/V_i(s)$ of the circuit in Fig. E.3(a). Show that the transfer function can be made independent of frequency if the condition $C_1R_1 = C_2R_2$ applies. Under this condition the circuit is called a *compensated attenuator*. Find the transmission of the compensated attenuator in terms of R_1 and R_2.

D **E.3 The circuit of Fig. E.3(a) is used as a compensated attenuator (see Problems E.1 and E.2) for an oscilloscope probe. The objective is to reduce the signal voltage applied to the input amplifier of the oscilloscope, with the signal attenuation independent of frequency. The probe itself includes R_1 and C_1, while R_2 and C_2 model the oscilloscope input circuit. For an oscilloscope having an input resistance of 1 MΩ and an input capacitance of 30 pF, design a compensated "10-to-1 probe" —that is, a probe that attenuates the input signal by a factor of 10. Find the input impedance of the probe when connected to the oscilloscope, which is the impedance seen by v_I in Fig. E.3(a). Show that this impedance is 10 times higher than that of the oscilloscope itself. This is the great advantage of the 10:1 probe.

E.4 In the circuits of Figs. E.4 and E.5, let $L = 10$ mH, $C = 0.01$ μF, and $R = 1$ kΩ. At what frequency does a phase angle of 45° occur?

***E.5** Consider a voltage amplifier with an open-circuit voltage gain $A_{vo} = -100$ V/V, $R_o = 0$, $R_i = 10$ kΩ, and

an input capacitance C_i (in parallel with R_i) of 10 pF. The amplifier has a feedback capacitance (a capacitance connected between output and input) $C_f = 1$ pF. The amplifier is fed with a voltage source V_s having a resistance $R_s = 10$ kΩ. Find the amplifier transfer function $V_o(s)/V_s(s)$ and sketch its magnitude response versus frequency (dB vs. frequency) on a log axis.

E.6 For the circuit in Fig. PE.6, assume the voltage amplifier to be ideal. Derive the transfer function $V_o(s)/V_i(s)$. What type of STC response is this? For $C = 0.01$ μF and $R = 100$ kΩ, find the corner frequency.

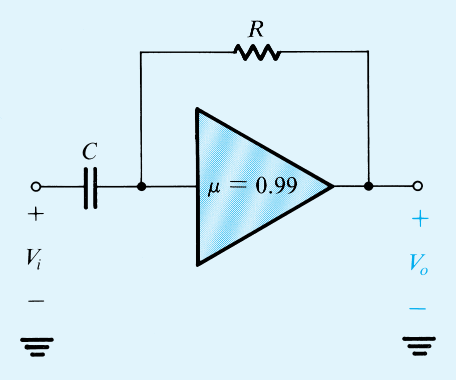


Figure PE.6

E.7 For the circuits of Figs. E.4(b) and E.5(b), find $v_O(t)$ if v_I is a 10-V step, $R = 1$ kΩ, and $L = 1$ mH.

E.8 Consider the exponential response of an STC low-pass circuit to a 10-V step input. In terms of the time constant τ, find the time taken for the output to reach 5 V, 9 V, 9.9 V, and 9.99 V.

E.9 The high-frequency response of an oscilloscope is specified to be like that of an STC LP circuit with a 100-MHz corner frequency. If this oscilloscope is used to display an ideal step waveform, what rise time (10% to 90%) would you expect to observe?

E.10 An oscilloscope whose step response is like that of a low-pass STC circuit has a rise time of t_s seconds. If an input signal having a rise time of t_w seconds is displayed, the waveform seen will have a rise time t_d seconds, which can be found using the empirical formula $t_d = \sqrt{t_s^2 + t_w^2}$. If $t_s = 35$ ns, what is the 3-dB frequency of the oscilloscope? What is the observed rise time for a waveform rising in 100 ns, 35 ns, and 10 ns? What is the actual rise time of a waveform whose displayed rise time is 49.5 ns?

E.11 A pulse of 10-ms width and 10-V amplitude is transmitted through a system characterized as having an STC high-pass response with a corner frequency of 10 Hz. What undershoot would you expect?

E.12 An RC differentiator having a time constant τ is used to implement a short-pulse detector. When a long pulse with $T \gg \tau$ is fed to the circuit, the positive and negative peak outputs are of equal magnitude. At what pulse width does the negative output peak differ from the positive one by 10%?

E.13 A high-pass STC circuit with a time constant of 1 ms is excited by a pulse of 10-V height and 1-ms width. Calculate the value of the undershoot in the output waveform. If an undershoot of 1 V or less is required, what is the time constant necessary?

E.14 A capacitor C is used to couple the output of an amplifier stage to the input of the next stage. If the first stage has an output resistance of 2 kΩ and the second stage has an input resistance of 3 kΩ, find the value of C so that a 1-ms pulse exhibits less than 1% sag. What is the associated 3-dB frequency?

D E.15 An RC differentiator is used to convert a step voltage change V to a single pulse for a digital-logic application. The logic circuit that the differentiator drives distinguishes signals above $V/2$ as "high" and below $V/2$ as "low." What must the time constant of the circuit be to convert a step input into a pulse that will be interpreted as "high" for 10 μs?

D E.16 Consider the circuit in Fig. E.7(a) with $\mu = -100$, $C_f = 100$ pF, and the amplifier being ideal. Find the value of R so that the gain $|V_o/V_s|$ has a 3-dB frequency of 1 kHz.

APPENDIX F

s-DOMAIN ANALYSIS: POLES, ZEROS, AND BODE PLOTS

In analyzing the frequency response of an amplifier, most of the work involves finding the amplifier voltage gain as a function of the complex frequency s. In this s-domain analysis, a capacitance C is replaced by an admittance sC, or equivalently an impedance $1/sC$, and an inductance L is replaced by an impedance sL. Then, using usual circuit-analysis techniques, one derives the voltage transfer function $T(s) \equiv V_o(s)/V_i(s)$.

EXERCISE

F.1 Find the voltage transfer function $T(s) \equiv V_o(s)/V_i(s)$ for the STC network shown in Fig. EF.1.

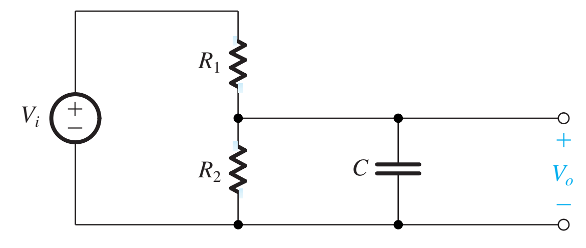


Figure EF.1

Ans. $T(s) = \dfrac{1/CR_1}{s + 1/C(R_1 \parallel R_2)}$

Once the transfer function $T(s)$ is obtained, it can be evaluated for **physical frequencies** by replacing s by $j\omega$. The resulting transfer function $T(j\omega)$ is in general a complex quantity whose magnitude gives the magnitude response (or transmission) and whose angle gives the phase response of the amplifier.

In many cases it will not be necessary to substitute $s = j\omega$ and evaluate $T(j\omega)$; rather, the form of $T(s)$ will reveal many useful facts about the circuit performance. In general, for all the circuits dealt with in this book, $T(s)$ can be expressed in the form

$$T(s) = \frac{a_m s^m + a_{m-1} s^{m-1} + \ldots + a_0}{s^n + b_{n-1} s^{n-1} + \ldots + b_0} \tag{F.1}$$

where the coefficients a and b are real numbers, and the order m of the numerator is smaller than or equal to the order n of the denominator; the latter is called the **order of the network**. Furthermore, for a **stable circuit**—that is, one that does not generate signals on its own—the denominator coefficients should be such that *the roots of the denominator polynomial all have negative real parts*. The problem of amplifier stability is studied in Chapter 10.

F.1 Poles and Zeros

An alternate form for expressing $T(s)$ is

$$T(s) = a_m \frac{(s - Z_1)(s - Z_2) \ldots (s - Z_m)}{(s - P_1)(s - P_2) \ldots (s - P_n)} \tag{F.2}$$

where a_m is a multiplicative constant (the coefficient of s^m in the numerator), $Z_1, Z_2, \ldots, Z_m$ are the roots of the numerator polynomial, and $P_1, P_2, \ldots, P_n$ are the roots of the denominator polynomial. $Z_1, Z_2, \ldots, Z_m$ are called the **transfer-function zeros** or **transmission zeros**, and $P_1, P_2, \ldots, P_n$ are the **transfer-function poles** or the **natural modes** of the network. A transfer function is completely specified in terms of its poles and zeros together with the value of the multiplicative constant.

The poles and zeros can be either real or complex numbers. However, since the a and b coefficients are real numbers, the complex poles (or zeros) must occur in **conjugate pairs**. That is, if $5 + j3$ is a zero, then $5 - j3$ also must be a zero. A zero that is purely imaginary ($\pm j\omega_Z$) causes the transfer function $T(j\omega)$ to be exactly zero at $\omega = \omega_Z$. This is because the numerator will have the factors $(s + j\omega_Z)(s - j\omega_Z) = (s^2 + \omega_Z^2)$, which for physical frequencies becomes $(-\omega^2 + \omega_Z^2)$, and thus the transfer fraction will be exactly zero at $\omega = \omega_Z$. Thus the "trap" one places at the input of a television set is a circuit that has a transmission zero at the particular interfering frequency. Real zeros, on the other hand, do not produce transmission nulls. Finally, note that for values of s much greater than all the poles and zeros, the transfer function in Eq. (F.1) becomes $T(s) \simeq a_m/s^{n-m}$. Thus the transfer function has $(n - m)$ zeros at $s = \infty$.

F.2 First-Order Functions

Many of the transfer functions encountered in this book have real poles and zeros and can therefore be written as the product of first-order transfer functions of the general form

$$T(s) = \frac{a_1 s + a_0}{s + \omega_0} \tag{F.3}$$

where $-\omega_0$ is the location of the real pole. The quantity ω_0, called the **pole frequency**, is equal to the inverse of the time constant of this single-time-constant (STC) network (see Appendix E). The constants a_0 and a_1 determine the type of STC network. Specifically, we studied in Chapter 1 two types of STC networks, low pass and high pass. For the low-pass first-order network we have

$$T(s) = \frac{a_0}{s + \omega_0} \tag{F.4}$$

In this case the dc gain is a_0/ω_0, and ω_0 is the corner or 3-dB frequency. Note that this transfer function has one zero at $s = \infty$. On the other hand, the first-order high-pass transfer function

has a zero at dc and can be written as

$$T(s) = \frac{a_1 s}{s + \omega_0} \tag{F.5}$$

At this point the reader is strongly urged to review the material on STC networks and their frequency and pulse responses in Appendix E. Of specific interest are the plots of the magnitude and phase responses of the two special kinds of STC networks. Such plots can be employed to generate the magnitude and phase plots of a high-order transfer function, as explained below.

F.3 Bode Plots

A simple technique exists for obtaining an approximate plot of the magnitude and phase of a transfer function given its poles and zeros. The technique is particularly useful in the case of real poles and zeros. The method was developed by H. Bode, and the resulting diagrams are called **Bode plots**.

A transfer function of the form depicted in Eq. (F.2) consists of a product of factors of the form $s + a$, where such a factor appears on top if it corresponds to a zero and on the bottom if it corresponds to a pole. It follows that the magnitude response in decibels of the network can be obtained by summing together terms of the form $20\log_{10}\sqrt{a^2+\omega^2}$, and the phase response can be obtained by summing terms of the form $\tan^{-1}(\omega/a)$. In both cases the terms corresponding to poles are summed with negative signs. For convenience we can extract the constant a and write the typical magnitude term in the form $20\log\sqrt{1+(\omega/a)^2}$. On a plot of decibels versus log frequency this term gives rise to the curve and straight-line asymptotes shown in Fig. F.1. Here the low-frequency asymptote is a horizontal straight line at 0-dB level and the high-frequency asymptote is a straight line with a slope of 6 dB/octave or, equivalently, 20 dB/decade. The two asymptotes meet at the frequency $\omega = |a|$, which is called the **corner frequency**. As indicated, the actual magnitude plot differs slightly from

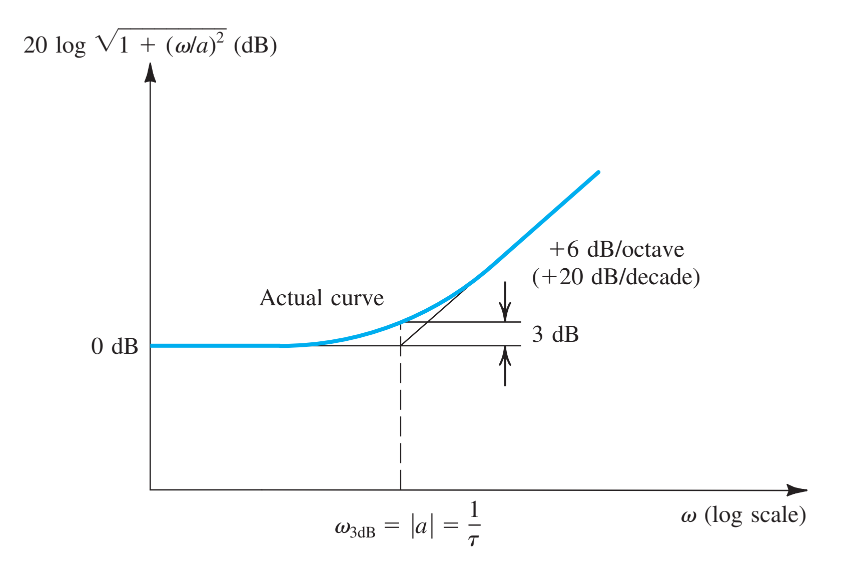


Figure F.1 Bode plot for the typical magnitude term. The curve shown applies for the case of a zero. For a pole, the high-frequency asymptote should be drawn with a –6-dB/octave slope.

the value given by the asymptotes; the maximum difference is 3 dB and occurs at the corner frequency.

For $a = 0$—that is, a pole or a zero at $s = 0$—the plot is simply a straight line of 6 dB/octave slope intersecting the 0-dB line at $\omega = 1$.

In summary, to obtain the Bode plot for the magnitude of a transfer function, the asymptotic plot for each pole and zero is first drawn. The slope of the high-frequency asymptote of the curve corresponding to a zero is +20 dB/decade, while that for a pole is −20 dB/decade. The various plots are then added together, and the overall curve is shifted vertically by an amount determined by the multiplicative constant of the transfer function.

Example F.1

An amplifier has the voltage transfer function

$$T(s) = \frac{10s}{\left(1 + s/10^2\right)\left(1 + s/10^5\right)}$$

Find the poles and zeros and sketch the magnitude of the gain versus frequency. Find approximate values for the gain at $\omega = 10$, 10^3, and 10^6 rad/s.

Solution

The zeros are as follows: one at $s = 0$ and one at $s = \infty$. The poles are as follows: one at $s = -10^2$ rad/s and one at $s = -10^5$ rad/s.

Figure F.2 shows the asymptotic Bode plots of the different factors of the transfer function. Curve 1, which is a straight line intersecting the ω-axis at 1 rad/s and having a +20 dB/decade slope, corresponds to the s term (that is, the zero at $s = 0$) in the numerator. The pole at $s = -10^2$ results in curve 2, which consists of two asymptotes intersecting at $\omega = 10^2$. Similarly, the pole at $s = -10^5$ is represented by curve 3, where the intersection of the asymptotes is at $\omega = 10^5$. Finally, curve 4 represents the multiplicative constant of value 10.

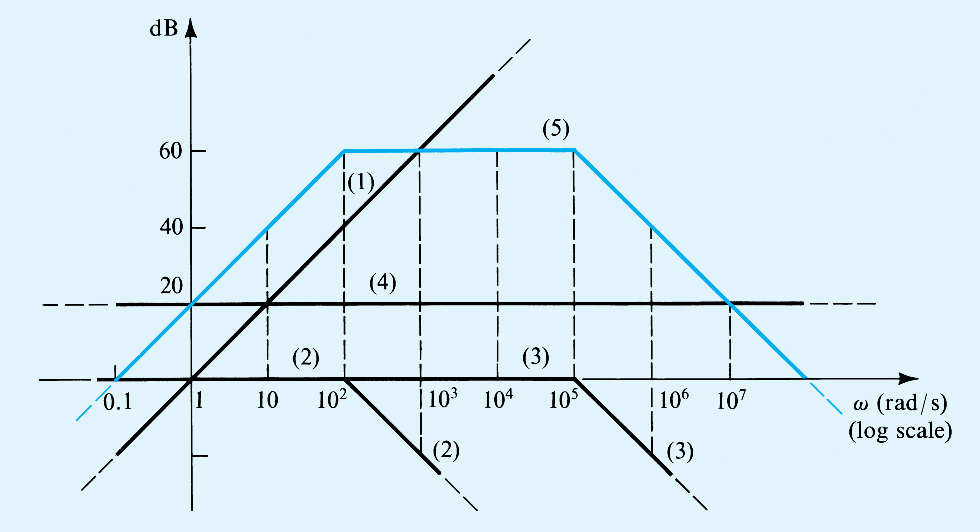


Figure F.2 Bode plots for Example F.1.

Adding the four curves results in the asymptotic Bode diagram of the amplifier gain (curve 5). Note that since the two poles are widely separated, the gain will be very close to 10^3 (60 dB) over the frequency range 10^2 to 10^5 rad/s. At the two corner frequencies (10^2 and 10^5 rad/s) the gain will be approximately 3 dB below the maximum of 60 dB. At the three specific frequencies, the values of the gain as obtained from the Bode plot and from exact evaluation of the transfer function are as follows:

ω	Approximate Gain	Exact Gain
10	40 dB	39.96 dB
10^3	60 dB	59.96 dB
10^6	40 dB	39.96 dB

We next consider the Bode phase plot. Figure F.3 shows a plot of the typical phase term $\tan^{-1}(\omega/a)$, assuming that a is negative. Also shown is an asymptotic straight-line approximation of the arctan function. The asymptotic plot consists of three straight lines. The first is horizontal at $\phi = 0$ and extends up to $\omega = 0.1|a|$. The second line has a slope of –45°/decade and extends from $\omega = 0.1|a|$ to $\omega = 10|a|$. The third line has a zero slope and a level of $\phi = -90°$. The complete phase response can be obtained by summing the asymptotic Bode plots of the phase of all poles and zeros.

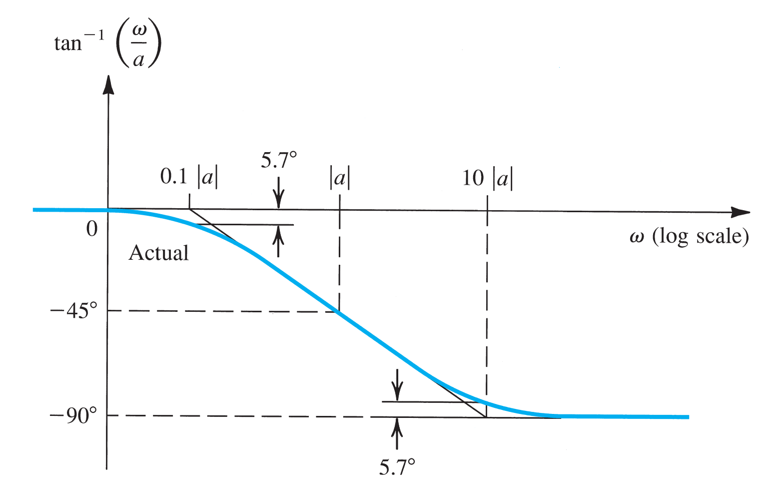


Figure F.3 Bode plot of the typical phase term $\tan^{-1}(\omega/a)$ when a is negative.

Example F.2

Find the Bode plot for the phase of the transfer function of the amplifier considered in Example F.1.

Solution

The zero at $s = 0$ gives rise to a constant +90° phase function represented by curve 1 in Fig. F.4. The pole at $s = -10^2$ gives rise to the phase function

$$\phi_1 = -\tan^{-1}\frac{\omega}{10^2}$$

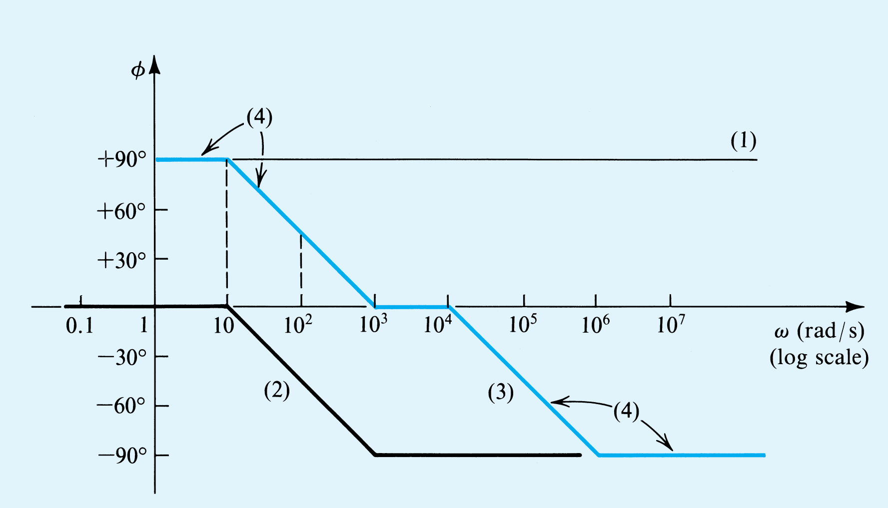


Figure F.4 Phase plots for Example F.2.

(the leading minus sign is due to the fact that this singularity is a pole). The asymptotic plot for this function is given by curve 2 in Fig. F.4. Similarly, the pole at $s = -10^5$ gives rise to the phase function

$$\phi_2 = -\tan^{-1}\frac{\omega}{10^5}$$

whose asymptotic plot is given by curve 3. The overall phase response (curve 4) is obtained by direct summation of the three plots. We see that at 100 rad/s, the amplifier phase leads by 45° and at 10^5 rad/s the phase lags by 45°.

F.4 An Important Remark

For constructing Bode plots, it is most convenient to express the transfer-function factors in the form $(1 + s/a)$. The material of Figs. F.1 and F.2 and of the preceding two examples is then directly applicable.

PROBLEMS

F.1 Find the transfer function $T(s) = V_o(s)/V_i(s)$ of the circuit in Fig. PF.1. Is this an STC network? If so, of what type? For $C_1 = C_2 = 0.5$ µF and $R = 100$ kΩ, find the location of the pole(s) and zero(s), and sketch Bode plots for the magnitude response and the phase response.

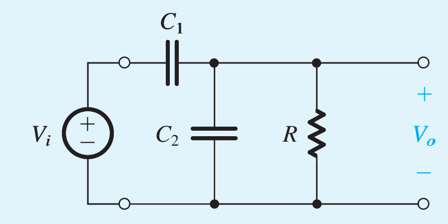


Figure PF.1

D *F.2 (a) Find the voltage transfer function $T(s) = V_o(s)/V_i(s)$, for the STC network shown in Fig. PF.2.

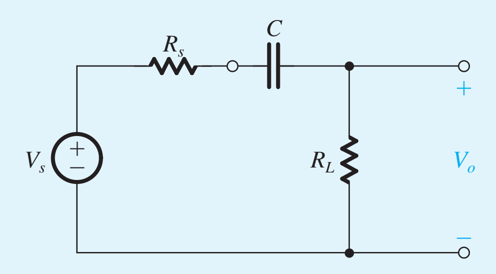


Figure PF.2

(b) In this circuit, capacitor C is used to couple the signal source V_s having a resistance R_s to a load R_L. For $R_s = 10\ \text{k}\Omega$, design the circuit, specifying the values of R_L and C to only one significant digit to meet the following requirements:

(i) The load resistance should be as small as possible.
(ii) The output signal should be at least 70% of the input at high frequencies.
(iii) The output should be at least 10% of the input at 10 Hz.

F.3 Two STC RC circuits, each with a pole at 100 rad/s and a maximum gain of unity, are connected in cascade with an intervening unity-gain buffer that ensures that they function separately. Characterize the possible combinations (of low-pass and high-pass circuits) by providing (i) the relevant transfer functions, (ii) the voltage gain at 10 rad/s, (iii) the voltage gain at 100 rad/s, and (iv) the voltage gain at 1000 rad/s.

F.4 Design the transfer function in Eq. (F.5) by specifying a_1 and ω_0 so that the gain is 10 V/V at high frequencies and 1 V/V at 10 Hz.

F.5 An amplifier has a low-pass STC frequency response. The magnitude of the gain is 20 dB at dc and 0 dB at 100 kHz. What is the corner frequency? At what frequency is the gain 19 dB? At what frequency is the phase $-6°$?

F.6 A transfer function has poles at (-5), $(-7 + j10)$, and (-20), and a zero at $(-1 - j20)$. Since this function represents an actual physical circuit, where must other poles and zeros be found?

F.7 An amplifier has a voltage transfer function $T(s) = 10^6 s(s+10)(s+10^3)$. Convert this to the form convenient for constructing Bode plots [that is, place the denominator factors in the form $(1 + s/a)$]. Provide a Bode plot for the magnitude response, and use it to find approximate values for the amplifier gain at 1, 10, 10^2, 10^3, 10^4, and 10^5 rad/s. What would the actual gain be at 10 rad/s? At 10^3 rad/s?

F.8 Find the Bode phase plot of the transfer function of the amplifier considered in Problem F.7. Estimate the phase angle at 1, 10, 10^2, 10^3, 10^4, and 10^5 rad/s. For comparison, calculate the actual phase at 1, 10, and 100 rad/s.

F.9 A transfer function has the following zeros and poles: one zero at $s = 0$ and one zero at $s = \infty$; one pole at $s = -100$ and one pole at $s = -10^6$. The magnitude of the transfer function at $\omega = 10^4$ rad/s is 100. Find the transfer function $T(s)$ and sketch a Bode plot for its magnitude.

F.10 Sketch Bode plots for the magnitude and phase of the transfer function

$$T(s) = \frac{10^4(1 + s/10^5)}{(1 + s/10^3)(1 + s/10^4)}$$

From your sketches, determine approximate values for the magnitude and phase at $\omega = 10^6$ rad/s. What are the exact values determined from the transfer function?

F.11 A particular amplifier has a voltage transfer function $T(s) = 10s^2/(1 + s/10)(1 + s/100)(1 + s/10^6)$. Find the poles and zeros. Sketch the magnitude of the gain in dB versus frequency on a logarithmic scale. Estimate the gain at 10^0, 10^3, 10^5, and 10^7 rad/s.

F.12 A direct-coupled differential amplifier has a differential gain of 100 V/V with poles at 10^6 and 10^8 rad/s, and a common-mode gain of 10^{-3} V/V with a zero at 10^4 rad/s and a pole at 10^8 rad/s. Sketch the Bode magnitude plots for the differential gain, the common-mode gain, and the CMRR. What is the CMRR at 10^7 rad/s? (*Hint:* Division of magnitudes corresponds to subtraction of logarithms.)

APPENDIX G

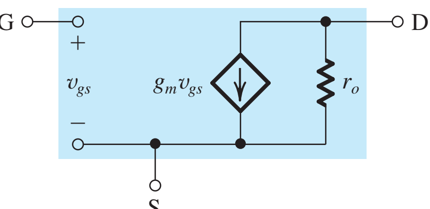
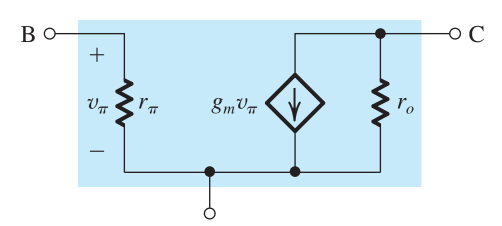
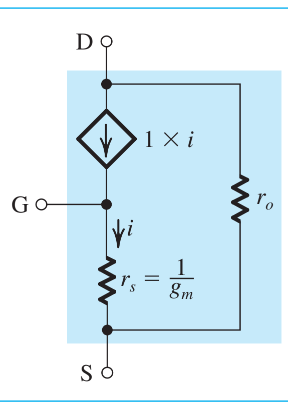
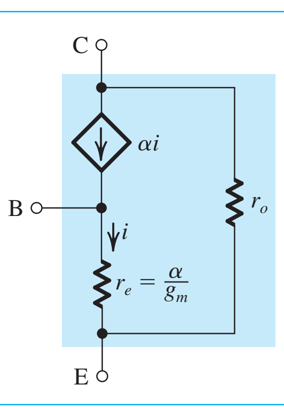
COMPARISON OF THE MOSFET AND THE BJT

The full text of Appendix G is on the book's website. Here, we show only Table G.3, which provides a summary comparison of the MOSFET and the BJT.

Table G.3 Comparison of the MOSFET and the BJT

	NMOS	*npn*
Circuit Symbol	i_D, i_G, v_{GD}, v_{DS}, v_{GS}	i_C, i_B, v_{BC}, v_{CE}, v_{BE}
To Operate in the Active Mode, Two Conditions Have to Be Satisfied	(1) *Induce a channel*: $v_{GS} \geq V_t, \quad V_t = 0.3\text{–}0.5\ \text{V}$ Let $v_{GS} = V_t + v_{OV}$ (2) *Pinch-off channel at drain*: $v_{GD} < V_t$ or equivalently, $v_{DS} \geq V_{OV}, \quad V_{OV} = 0.1\text{–}0.3\ \text{V}$	(1) *Forward-bias EBJ:* $v_{BE} \geq V_{BE\text{on}}, \quad V_{BE\text{on}} \simeq 0.5\ \text{V}$ (2) *Reverse-bias CBJ:* $v_{BC} < V_{BC\text{on}}, \quad V_{BC\text{on}} \simeq 0.4\ \text{V}$ or equivalently, $v_{CE} \geq 0.3\ \text{V}$
Current–Voltage Characteristics in the Active Region	$i_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(v_{GS} - V_t)^2\left(1 + \frac{v_{DS}}{V_A}\right)$ $= \frac{1}{2}\mu_n C_{ox}\frac{W}{L}v_{OV}^2\left(1 + \frac{v_{DS}}{V_A}\right)$ $i_G = 0$	$i_C = I_S e^{v_{BE}/V_T}\left(1 + \frac{v_{CE}}{V_A}\right)$ $i_B = i_C/\beta$

Table G.3 *continued*

	NMOS	*npn*
Low-Frequency, Hybrid-π Model		
Low-Frequency T Model		
Transconductance g_m	$g_m = I_D/(V_{OV}/2)$ $g_m = (\mu_n C_{ox})\left(\frac{W}{L}\right)V_{OV}$ $g_m = \sqrt{2(\mu_n C_{ox})\left(\frac{W}{L}\right)I_D}$	$g_m = I_C/V_T$
Output Resistance r_o	$r_o = V_A/I_D = \frac{V_A'L}{I_D}$	$r_o = V_A/I_C$
Intrinsic Gain $A_0 \equiv g_m r_o$	$A_0 = V_A/(V_{OV}/2)$ $A_0 = \frac{2V_A'L}{V_{OV}}$ $A_0 = \frac{V_A'\sqrt{2\mu_n C_{ox} WL}}{\sqrt{I_D}}$	$A_0 = V_A/V_T$
Input Resistance with Source (Emitter) Grounded	∞	$r_\pi = \beta/g_m$

(*continued*)

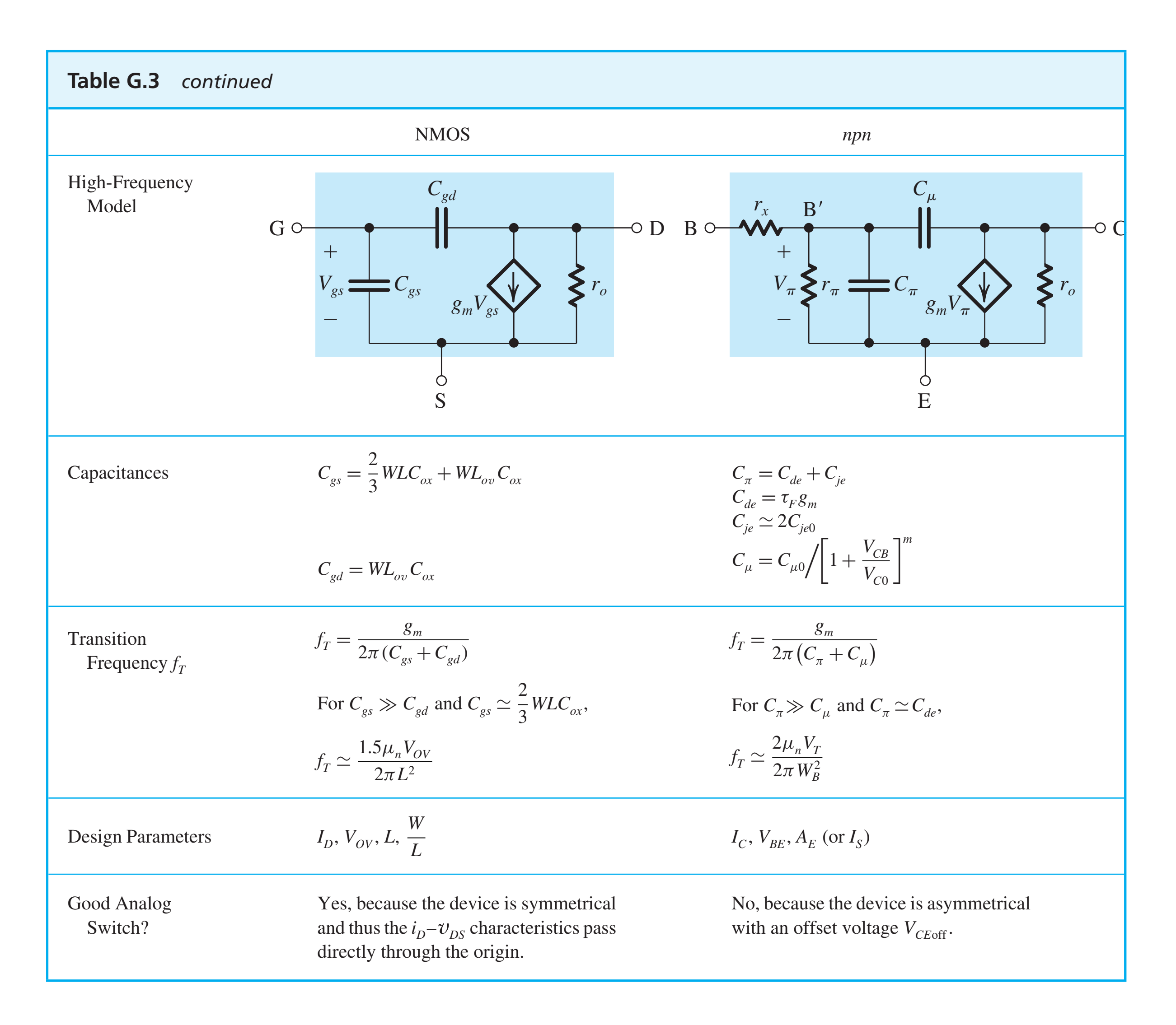


Table G.3 *continued*

	NMOS	*npn*
High-Frequency Model	(circuit: NMOS high-frequency model)	(circuit: npn high-frequency model)
Capacitances	$C_{gs} = \frac{2}{3}WLC_{ox} + WL_{ov}C_{ox}$ $C_{gd} = WL_{ov}C_{ox}$	$C_\pi = C_{de} + C_{je}$ $C_{de} = \tau_F g_m$ $C_{je} \simeq 2C_{je0}$ $C_\mu = C_{\mu 0} \Big/ \left[1 + \frac{V_{CB}}{V_{C0}}\right]^m$
Transition Frequency f_T	$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$ For $C_{gs} \gg C_{gd}$ and $C_{gs} \simeq \frac{2}{3}WLC_{ox}$, $f_T \simeq \frac{1.5\mu_n V_{OV}}{2\pi L^2}$	$f_T = \frac{g_m}{2\pi(C_\pi + C_\mu)}$ For $C_\pi \gg C_\mu$ and $C_\pi \simeq C_{de}$, $f_T \simeq \frac{2\mu_n V_T}{2\pi W_B^2}$
Design Parameters	$I_D, V_{OV}, L, \frac{W}{L}$	I_C, V_{BE}, A_E (or I_S)
Good Analog Switch?	Yes, because the device is symmetrical and thus the i_D–v_{DS} characteristics pass directly through the origin.	No, because the device is asymmetrical with an offset voltage V_{CEoff}.

APPENDIX G

COMPARISON OF THE MOSFET AND THE BJT

In this appendix we present a comparison of the characteristics of the two major electronic devices: the MOSFET and the BJT. To facilitate this comparison, typical values for the important parameters of the two devices are first presented. We also discuss the design parameters available with each of the two devices, such as I_C in the BJT, and I_D and V_{OV} in the MOSFET, and the trade-offs encountered in deciding on suitable values for these.

G.1 Typical Values of MOSFET Parameters

Typical values for the important parameters of NMOS and PMOS transistors fabricated in a number of CMOS processes are shown in Table G.1. Each process is characterized by the minimum allowed channel length, L_{min}; thus, for example, in a 0.18-μm process, the smallest transistor has a channel length $L = 0.18\ \mu\text{m}$. The technologies presented in Table G.1 are in descending order of channel length, with that having the shortest channel length being the most modern. Although the 0.8-μm process is now obsolete, its data are included to show trends in the values of various parameters. It should also be mentioned that although Table G.1 stops at the 65-nm process, by 2014 there were 45-, 32-, and 22-nm processes available, and processes down to 14 nm were in various stages of development. The 0.18-μm and the 0.13-μm processes, however, remained popular in the design of analog ICs. The most recently announced digital ICs utilize 32-nm and 22-nm processes and pack as many as 4.3 billion transistors onto one chip. An important caution is in order regarding the data presented in Table G.1: These data do *not* pertain to any particular commercially available process.

Table G.1 Typical Values of CMOS Device Parameters

	0.8 nm		0.5 nm		0.25 nm		0.18 nm		0.13 nm		65 nm	
Parameter	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS
t_{ox} (nm)	15	15	9	9	6	6	4	4	2.7	2.7	1.4	1.4
C_{ox} (fF/μm^2)	2.3	2.3	3.8	3.8	5.8	5.8	8.6	8.6	12.8	12.8	25	25
μ (cm^2/V·s)	550	250	500	180	460	160	450	100	400	100	216	40
μC_{ox} (μA/V^2)	127	58	190	68	267	93	387	86	511	128	540	100
V_{t0} (V)	0.7	−0.7	0.7	−0.8	0.5	−0.6	0.5	−0.5	0.4	−0.4	0.35	−0.35
V_{DD} (V)	5	5	3.3	3.3	2.5	2.5	1.8	1.8	1.3	1.3	1.0	1.0
$\lvert V_A' \rvert$ (V/μm)	25	20	20	10	5	6	5	6	5	6	3	3
C_{ov} (fF/μm)	0.2	0.2	0.4	0.4	0.3	0.3	0.37	0.33	0.36	0.33	0.33	0.31

Accordingly, these generic data are not intended for use in an actual IC design; rather, they show trends and, as we shall see, help to illustrate design trade-offs as well as enable us to work out design examples and problems with parameter values that are as realistic as possible.

As indicated in Table G.1, the trend has been to reduce the minimum allowable channel length. This trend has been motivated by the desire to pack more transistors on a chip as well as to operate at higher speeds or, in analog terms, over wider bandwidths.

Observe that the oxide thickness, t_{ox}, scales down with the channel length, reaching 1.4 nm for the 65-nm process. Since the oxide capacitance C_{ox} is inversely proportional to t_{ox}, we see that C_{ox} increases as the technology scales down. The surface mobility μ decreases as the technology minimum-feature size is decreased, and μ_p decreases faster than μ_n. As a result, the ratio of μ_p to μ_n has been decreasing with each generation of technology, falling from about 0.5 for older technologies to 0.2 or so for the newer ones. Despite the reduction of μ_n and μ_p, the transconductance parameters $k_n' = \mu_n C_{ox}$ and $k_p' = \mu_p C_{ox}$ have been steadily increasing. As a result, modern short-channel devices achieve required levels of bias currents at lower overdrive voltages. As well, they achieve higher transconductance, a major advantage.

Although the magnitudes of the threshold voltages V_{tn} and V_{tp} have been decreasing with $L_{\min}$ from about 0.7–0.8 V to 0.3–0.4 V, the reduction has not been as large as that of the power supply V_{DD}. The latter has been reduced dramatically, from 5 V for older technologies to 1.0 V for the 65-nm process. This reduction has been necessitated by the need to keep the electric fields in the smaller devices from reaching very high values. Another reason for reducing V_{DD} is to keep power dissipation as low as possible given that the IC chip now has a much larger number of transistors.[1]

The fact that in modern short-channel CMOS processes $|V_t|$ has become a much larger proportion of the power-supply voltage poses a serious challenge to the circuit design engineer. Recalling that $|V_{GS}| = |V_t| + |V_{OV}|$, where V_{OV} is the overdrive voltage, to keep $|V_{GS}|$ reasonably small, $|V_{OV}|$ for modern technologies is usually in the range of 0.1 V to 0.2 V. To appreciate this point further, recall that to operate a MOSFET in the saturation region, $|V_{DS}|$ must exceed $|V_{OV}|$; thus, to be able to have a number of devices stacked between the power-supply rails in a regime in which V_{DD} is only 1.8 V or lower, we need to keep $|V_{OV}|$ as low as possible. We will shortly see, however, that operating at a low $|V_{OV}|$ has some drawbacks.

Another significant though undesirable feature of modern deep submicron ($L_{\min} <$ 0.25 μm) CMOS technologies is that the channel-length modulation effect is very pronounced. As a result, V_A' has decreased to about 3 V/μm, which combined with the decreasing values of L has caused the Early voltage $V_A = V_A' L$ to become very small. Correspondingly, short-channel MOSFETs exhibit low output resistances.

Studying the MOSFET high-frequency[2] equivalent-circuit model in Section 10.2 and the high-frequency response of the common-source amplifier in Section 10.3 shows that two major MOSFET capacitances are C_{gs} and C_{gd}. While C_{gs} has an overlap component,[3] C_{gd} is entirely an overlap capacitance. Both C_{gd} and the overlap component of C_{gs} are almost equal and are denoted C_{ov}. The last line of Table G.1 provides the value of C_{ov} per micron of gate width. Although the normalized C_{ov} has been staying more or less constant with the reduction in $L_{\min}$, we will shortly see that the shorter devices exhibit much higher operating speeds and wider amplifier bandwidths than the longer devices. Specifically, we will, for example, see that f_T for a 0.25-μm NMOS transistor can be as high as 10 GHz.

[1] Chip power dissipation is a very serious issue, with some ICs dissipating as much as 100 W. As a result, an important current area of research concerns what is termed "power-aware design."

[2] For completeness, this appendix includes material on the high-frequency models and operation of both the MOSFET and the BJT. These topics are covered in Chapter 10. The reader can easily skip the appendix paragraphs dealing with these topics until Chapter 10 has been studied.

[3] Overlap capacitances result because the gate electrode overlaps the source and drain diffusions (Fig. 5.1).

G.2 Typical Values of IC BJT Parameters

Table G.2 provides typical values for the major parameters that characterize integrated-circuit bipolar transistors. Data are provided for devices fabricated in two different processes: the standard, old process, known as the "high-voltage process," and an advanced, modern process, referred to as a "low-voltage process." For each process we show the parameters of the standard *npn* transistor and those of a special type of *pnp* transistor known as a **lateral *pnp*** (as opposed to **vertical**, as in the *npn* case) (see Appendix A). In this regard we should mention that a major drawback of standard bipolar integrated-circuit fabrication processes has been the lack of *pnp* transistors of a quality equal to that of the *npn* devices. Rather, there are a number of *pnp* implementations for which the lateral *pnp* is the most economical to fabricate. Unfortunately, however, as should be evident from Table G.2, the lateral *pnp* has characteristics that are much inferior to those of the vertical *npn*. Note in particular the lower value of β and the much larger value of the forward transit time τ_F that determines the emitter–base diffusion capacitance C_{de} and, hence, the transistor speed of operation. The data in Table G.2 can be used to show that the unity-gain frequency of the lateral *pnp* is 2 orders of magnitude lower than that of the *npn* transistor fabricated in the same process. Another important difference between the lateral *pnp* and the corresponding *npn* transistor is the value of collector current at which their β values reach their maximums: For the high-voltage process, for example, this current is in the tens of microamperes range for the *pnp* and in the milliampere range for the *npn*. On the positive side, the problem of the lack of high-quality *pnp* transistors has spurred analog circuit designers to come up with highly innovative circuit topologies that either minimize the use of *pnp* transistors or minimize the dependence of circuit performance on that of the *pnp*. We encounter some of these ingenious circuits at various locations in this book.

The dramatic reduction in device size achieved in the advanced low-voltage process should be evident from Table G.2. As a result, the scale current I_S also has been reduced by about three orders of magnitude. Here we should note that the base width, W_B, achieved in the advanced process is on the order of 0.1 μm, as compared to a few microns in the standard high-voltage process. Note also the dramatic increase in speed; for the low-voltage *npn* transistor, $\tau_F = 10$ ps as opposed to 0.35 ns in the high-voltage process. As a result, f_T for the modern *npn* transistor is 10 GHz to 25 GHz, as compared to the 400 MHz to 600 MHz achieved in the high-voltage process. Although the Early voltage, V_A, for the modern process is lower than its value in the

Table G.2 Typical Parameter Values for BJTs*

	Standard High-Voltage Process		Advanced Low-Voltage Process	
Parameter	*npn*	Lateral *pnp*	*npn*	Lateral *pnp*
A_E (μm^2)	500	900	2	2
I_S (A)	5×10^{-15}	2×10^{-15}	6×10^{-18}	6×10^{-18}
β_0 (A/A)	200	50	100	50
V_A (V)	130	50	35	30
V_{CEO} (V)	50	60	8	18
τ_F	0.35 ns	30 ns	10 ps	650 ps
C_{je0}	1 pF	0.3 PF	5 fF	14 fF
$C_{\mu 0}$	0.3 pF	1 pF	5 fF	15 fF
r_x (Ω)	200	300	400	200

*Adapted from Gray et al. (2001); see Appendix I.

old high-voltage process, it is still reasonably high at 35 V. Another feature of the advanced process—and one that is not obvious from Table G.2—is that β for the *npn* peaks at a collector current of 50 μA or so. Finally, note that as the name implies, *npn* transistors fabricated in the low-voltage process break down at collector–emitter voltages of 8 V, versus 50 V or so for the high-voltage process. Thus, while circuits designed with the standard high-voltage process utilize power supplies of ±15 V (e.g., in commercially available op amps of the 741 type), the total power-supply voltage utilized with modern bipolar devices is 5 V (or even 2.5 V to achieve compatibility with some of the submicron CMOS processes).

G.3 Comparison of Important Characteristics

Table G.3 provides a compilation of the important characteristics of the NMOS and the *npn* transistors. The material is presented in a manner that facilitates comparison. In the following, we comment on the various items in Table G.3. As well, a number of numerical examples and exercises are provided to illustrate how the wealth of information in Table G.3 can be put to use. Before proceeding, note that the PMOS and the *pnp* transistors can be compared in a similar way.

Table G.3 Comparison of MOSFET and the BJT

	NMOS	*npn*
Circuit Symbol	[NMOS symbol: i_D, i_G, v_{GD}, v_{DS}, v_{GS}]	[npn symbol: i_C, i_B, v_{BC}, v_{CE}, v_{BE}]
To Operate in the Active Mode, Two Conditions Have to Be Satisfied	(1) *Induce a channel*: $v_{GS} \geq V_t, \quad V_t = 0.3\text{–}0.5\text{ V}$ Let $v_{GS} = V_t + v_{OV}$ (2) *Pinch-off channel at drain*: $v_{GD} < V_t$ or equivalently, $v_{DS} \geq V_{OV}, \quad V_{OV} = 0.1\text{–}0.3\text{ V}$	(1) *Forward-bias EBJ:* $v_{BE} \geq V_{BEon}, \quad V_{BEon} \simeq 0.5\text{ V}$ (2) *Reverse-bias CBJ:* $v_{BC} < V_{BCon}, \quad V_{BCon} \simeq 0.4\text{ V}$ or equivalently, $v_{CE} \geq 0.3\text{ V}$
Current–Voltage Characteristics in the Active Region	$i_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(v_{GS} - V_t)^2\left(1 + \frac{v_{DS}}{V_A}\right)$ $= \frac{1}{2}\mu_n C_{ox}\frac{W}{L}v_{OV}^2\left(1 + \frac{v_{DS}}{V_A}\right)$ $i_G = 0$	$i_C = I_S e^{v_{BE}/V_T}\left(1 + \frac{v_{CE}}{V_A}\right)$ $i_B = i_C/\beta$

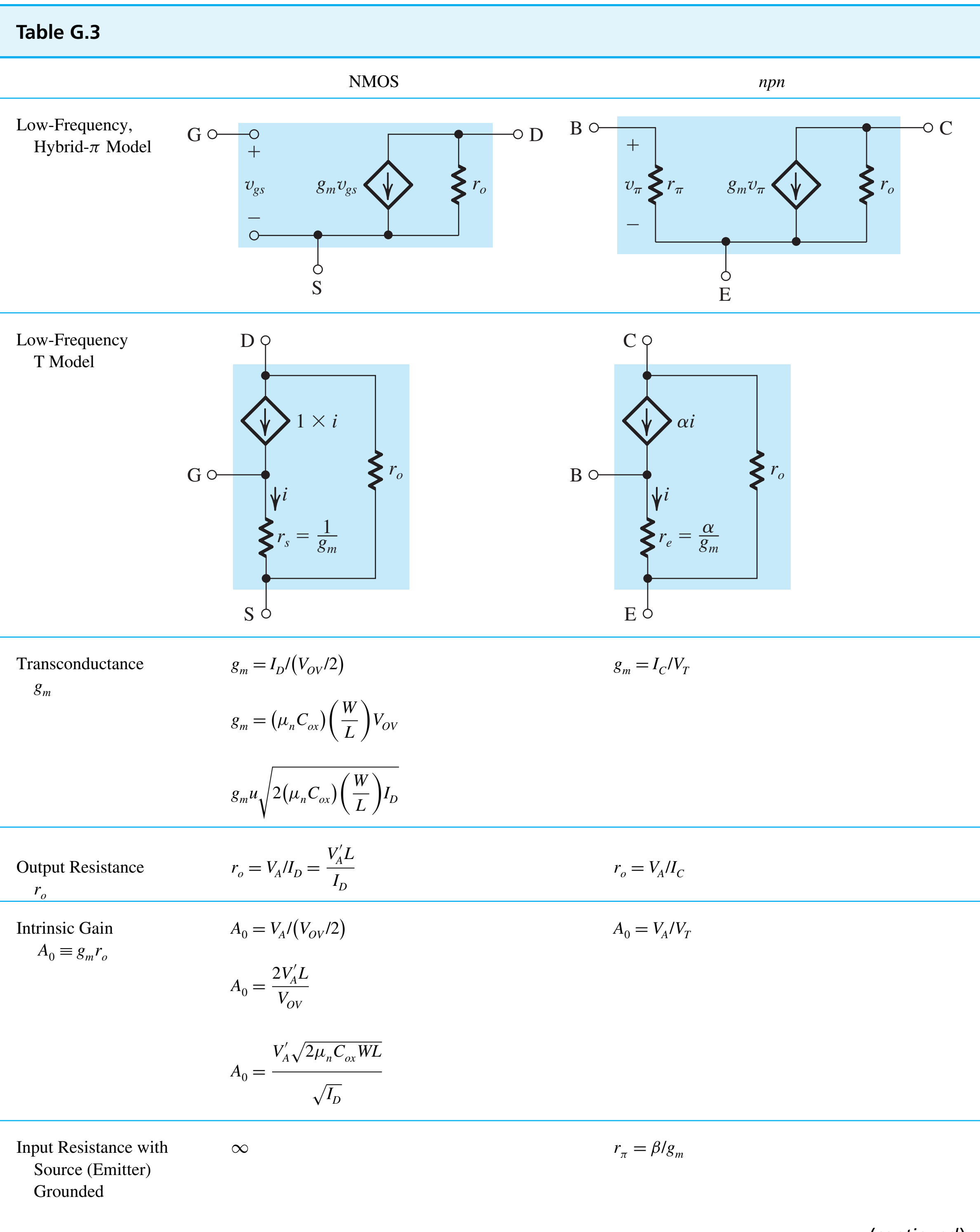


Table G.3

	NMOS	*npn*
Low-Frequency, Hybrid-π Model		
Low-Frequency T Model		
Transconductance g_m	$g_m = I_D/(V_{OV}/2)$ $g_m = (\mu_n C_{ox})\left(\frac{W}{L}\right)V_{OV}$ $g_m u\sqrt{2(\mu_n C_{ox})\left(\frac{W}{L}\right)I_D}$	$g_m = I_C/V_T$
Output Resistance r_o	$r_o = V_A/I_D = \frac{V_A'L}{I_D}$	$r_o = V_A/I_C$
Intrinsic Gain $A_0 \equiv g_m r_o$	$A_0 = V_A/(V_{OV}/2)$ $A_0 = \frac{2V_A'L}{V_{OV}}$ $A_0 = \frac{V_A'\sqrt{2\mu_n C_{ox} WL}}{\sqrt{I_D}}$	$A_0 = V_A/V_T$
Input Resistance with Source (Emitter) Grounded	∞	$r_\pi = \beta/g_m$

(*continued*)

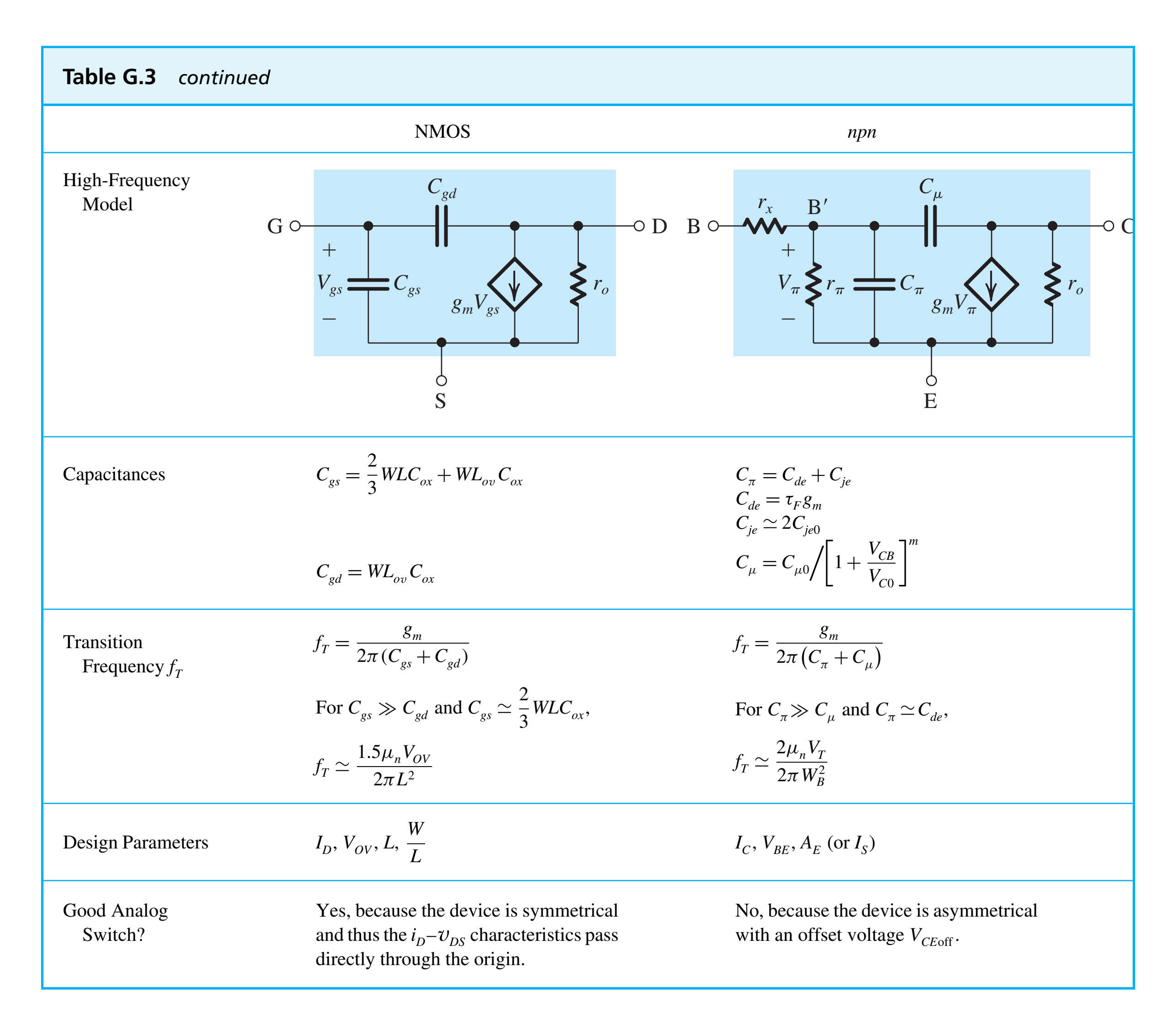


Table G.3 *continued*

	NMOS	*npn*
High-Frequency Model		
Capacitances	$C_{gs} = \frac{2}{3}WLC_{ox} + WL_{ov}C_{ox}$ $C_{gd} = WL_{ov}C_{ox}$	$C_\pi = C_{de} + C_{je}$ $C_{de} = \tau_F g_m$ $C_{je} \simeq 2C_{je0}$ $C_\mu = C_{\mu 0} \Big/ \left[1 + \frac{V_{CB}}{V_{C0}}\right]^m$
Transition Frequency f_T	$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$ For $C_{gs} \gg C_{gd}$ and $C_{gs} \simeq \frac{2}{3}WLC_{ox}$, $f_T \simeq \frac{1.5\mu_n V_{OV}}{2\pi L^2}$	$f_T = \frac{g_m}{2\pi(C_\pi + C_\mu)}$ For $C_\pi \gg C_\mu$ and $C_\pi \simeq C_{de}$, $f_T \simeq \frac{2\mu_n V_T}{2\pi W_B^2}$
Design Parameters	$I_D, V_{OV}, L, \frac{W}{L}$	I_C, V_{BE}, A_E (or I_S)
Good Analog Switch?	Yes, because the device is symmetrical and thus the i_D–v_{DS} characteristics pass directly through the origin.	No, because the device is asymmetrical with an offset voltage V_{CEoff}.

G.3.1 Operating Conditions

At the outset, note that we shall use **active mode** or **active region** to denote both the active mode of operation of the BJT and the saturation mode of operation of the MOSFET.

The conditions for operating in the active mode are very similar for the two devices: The explicit threshold V_t of the MOSFET has V_{BEon} as its implicit counterpart in the BJT. Furthermore, for modern processes, V_{BEon} and V_t are almost equal.

Also, pinching off the channel of the MOSFET at the drain end is very similar to reverse biasing the CBJ of the BJT; the first makes i_D nearly independent of v_D, and the second makes I_C nearly independent of v_C. Note, however, that the asymmetry of the BJT results in V_{BCon} and V_{BEon} being unequal, while in the symmetrical MOSFET the operative threshold voltages at the source and the drain ends of the channel are identical (V_t). Finally, for both the MOSFET and the BJT to operate in the active mode, the voltage across the device (v_{DS}, v_{CE}) must be at least 0.1 V to 0.3 V.

G.3.2 Current–Voltage Characteristics

The square-law control characteristic, $i_D - v_{GS}$, in the MOSFET should be contrasted with the exponential control characteristic, $i_C - v_{BE}$, of the BJT. Obviously, the latter is a much more sensitive relationship, with the result that i_C can vary over a very wide range (five decades or more) within the same BJT. In the MOSFET, the range of i_D achieved in the same device is much more limited. To appreciate this point further, consider the parabolic relationship between i_D and v_{OV}, and recall from our discussion above that v_{OV} is usually kept in a narrow range (0.1 V to 0.3 V).

Next we consider the effect of the device dimensions on its current. For the bipolar transistor, the control parameter is the area of the emitter–base junction (EBJ), A_E, which determines the scale current I_S. It can be varied over a relatively narrow range, such as 10 to 1. Thus, while the emitter area can be used to achieve current scaling in an I_C (as we can see in Section 8.2 in connection with the design of current mirrors), its narrow range of variation reduces its significance as a design parameter. This is particularly so if we compare A_E with its counterpart in the MOSFET, the aspect ratio W/L. MOSFET devices can be designed with W/L ratios in a wide range, such as 1.0 to 500. As a result, W/L is a very significant MOS design parameter. Like A_E, it is also used in current scaling, as we can see in Section 8.2. Combining the possible range of variation of v_{OV} and W/L, one can design MOS transistors to operate over an i_D range of four decades or so.

The channel-length modulation in the MOSFET and the base-width modulation in the BJT are similarly modeled and give rise to the dependence of i_D (i_C) on v_{DS} (v_{CE}) and, hence, to the finite output resistance r_o in the active region. Two important differences, however, exist. In the BJT, V_A is solely a process-technology parameter and does not depend on the dimensions of the BJT. In the MOSFET, the situation is quite different: $V_A = V_A'L$, where V_A' is a process-technology parameter and L is the channel length used. Also, in modern deep submicron processes, V_A' is very low, resulting in V_A values that are lower than the corresponding values for the BJT.

The last, and perhaps most important, difference between the current–voltage characteristics of the two devices concerns the input current into the control terminal: While at low frequencies the gate current of the MOSFET is practically zero and the input resistance looking into the gate is practically infinite, the BJT draws base current i_B that is proportional to the collector current; that is, $i_B = i_C/\beta$ The finite base current and the corresponding finite input resistance looking into the base comprise a definite disadvantage of the BJT in comparison to the MOSFET. Indeed, it is the infinite input resistance of the MOSFET that has made possible analog and digital circuit applications that are not feasible with the BJT. Examples include dynamic digital memory (Chapter 16) and switched-capacitor filters (Chapter 17).

Example G.1

(a) For an NMOS transistor with $W/L = 10$ fabricated in the 0.18-μm process whose data are given in Table G.1, find the values of V_{OV} and V_{GS} required to operate the device at $I_D = 100$ μA. Ignore channel-length modulation.

(b) Find V_{BE} for an *npn* transistor fabricated in the low-voltage process specified in Table G.2 and operated at $I_C = 100$ μA. Ignore base-width modulation.

Example G.1 *continued*

Solution

(a)

$$I_D = \frac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)V_{OV}^2$$

Substituting $I_D = 100\ \mu A$, $W/L = 10$, and, from Table G.1, $\mu_n C_{ox} = 387\ \mu A/V^2$ results in

$$100 = \frac{1}{2} \times 387 \times 10 \times V_{OV}^2$$

$$V_{OV} = 0.23\text{ V}$$

Thus,

$$V_{GS} = V_{tn} + V_{OV} = 0.5 + 0.23 = 0.73\text{ V}$$

(b)

$$I_C = I_S e^{V_{BE}/V_T}$$

Substituting $I_C = 100\ \mu A$ and, from Table G.2, $I_S = 6 \times 10^{-18}$ A gives,

$$V_{BE} = 0.025\ \ln\frac{100 \times 10^{-6}}{6 \times 10^{-18}} = 0.76\text{ V}$$

EXERCISE

G.1 (a) For NMOS transistors fabricated in the 0.18-µm technology specified in Table G.1, find the range of I_D obtained for active-mode operation with V_{OV} ranging from 0.2 V to 0.4 V and $W/L = 0.1$ to 100. Neglect channel-length modulation.
(b) If a similar range of current is required in an *npn* transistor fabricated in the low-voltage process specified in Table G.2, find the corresponding change in its V_{BE}.
Ans. (a) $I_{Dmin} = 0.8\ \mu A$ and $I_{Dmax} = 3.1$ mA for a range of about 4000:1; (b) for I_C varying over a 4000:1 range, $\Delta V_{BE} = 207$ mV

G.3.3 Low-Frequency Small-Signal Models

The low-frequency models for the two devices are very similar except, of course, for the finite base current (finite β) of the BJT, which gives rise to r_π in the hybrid-π model and to the unequal emitter and collector currents in the T models $\alpha < 1$. Here it is interesting to note that the low-frequency, small-signal models become identical if one thinks of the MOSFET as a BJT with $\beta = \infty(\alpha = 1)$

For both devices, the hybrid-π model indicates that the **open-circuit voltage gain** obtained from gate to drain (base to collector) with the source (emitter) grounded is $-g_m r_o$. It follows that $g_m r_o$ is the *maximum gain available from a single transistor* of either type. This important transistor parameter is given the name **intrinsic gain** and is denoted A_0. We have more to say about the intrinsic gain in Section 8.3.2.

Although not included in the MOSFET low-frequency model shown in Table G.3, the body effect can have some implications for the operation of the MOSFET as an amplifier. In simple terms, if the body (substrate) is not connected to the source, it can act as a second gate for the MOSFET. The voltage signal that develops between the body and the source, v_{bs}, gives rise to a drain current component $g_{mb}v_{bs}$, where the body transconductance g_{mb} is proportional to g_m; that is, $g_{mb} = \chi g_m$, where the factor χ is in the range of 0.1 to 0.2. The body effect has no counterpart in the BJT.

G.3.4 The Transconductance

For the BJT, the transconductance g_m depends *only* on the dc collector current I_C. (Recall that V_T is a physical constant $\simeq$ 0.025 V at room temperature.) It is interesting to observe that g_m does not depend on the geometry of the BJT, and its dependence on the EBJ area is only through the effect of the area on the total collector current I_C. Similarly, the dependence of g_m on V_{BE} is only through the fact that V_{BE} determines the total current in the collector. By contrast, g_m of the MOSFET depends on I_D, V_{OV}, and W/L. Therefore, we use three different (but equivalent) formulas to express g_m of the MOSFET.

The first formula given in Table G.3 for the MOSFET's g_m is the most directly comparable with the formula for the BJT. It indicates that for the same operating current, g_m of the MOSFET is smaller than that of the BJT. This is because $V_{OV}/2$ is the range of 0.05 V to 0.15 V, which is two to six times the corresponding term in the BJT's formula, namely V_T.

The second formula for the MOSFET's g_m indicates that for a given device (i.e., given W/L), g_m is proportional to V_{OV}. Thus a higher g_m is obtained by operating the MOSFET at a higher overdrive voltage. However, we should recall the limitations imposed on the magnitude of V_{OV} by the limited value of V_{DD}. Put differently, the need to obtain a reasonably high g_m constrains the designer's interest in reducing V_{OV}.

The third g_m formula shows that for a given transistor (i.e., given W/L), g_m is proportional to $\sqrt{I_D}$. This should be contrasted with the bipolar case, where g_m is directly proportional to I_C.

G.3.5 Output Resistance

The output resistance for both devices is determined by similar formulas, with r_o being the ratio of V_A to the bias current (I_D or I_C). Thus, for both transistors, r_o is inversely proportional to the bias current. The difference in nature and magnitude of V_A between the two devices has already been discussed.

G.3.6 Intrinsic Gain

The intrinsic gain A_0 of the BJT is the ratio of V_A, which is solely a process parameter (5 V to 100 V), and V_T, which is a physical parameter (0.025 V at room temperature). Thus A_0 of a BJT is independent of the device junction area and of the operating current, and its value ranges from 200 V/V to 5000 V/V. The situation in the MOSFET is very different: Table G.3 provides three different (but equivalent) formulas for expressing the MOSFET's intrinsic

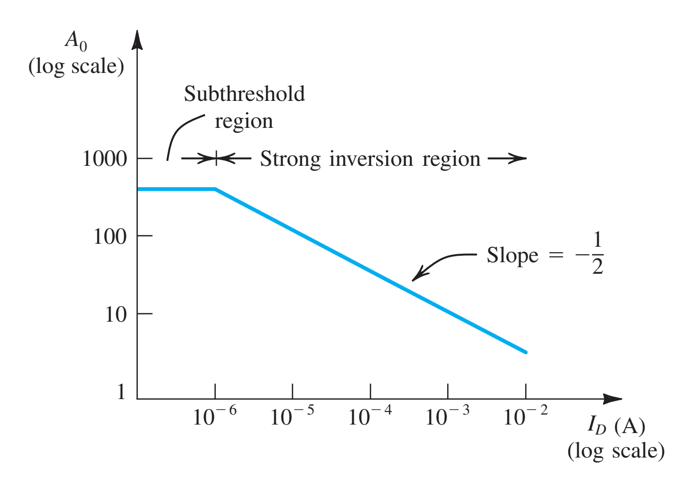


Figure G.1 The intrinsic gain of the MOSFET versus bias current I_D. Outside the subthreshold region, this is a plot of $A_0 = V_A'\sqrt{2\mu_n C_{ox} WL/I_D}$, for the case: $\mu_n C_{ox} = 20\ \mu\text{A/V}^2$, $V_A' = 20$ V/µm, $L = 2$ µm, and $W = 20$ µm.

gain. The first formula is the one most directly comparable to that of the BJT. Here, however, we note the following:

1. The quantity in the denominator is $V_{OV}/2$, which is a design parameter, and although it is becoming smaller in designs using short-channel technologies, it is still at least two to four times larger than V_T. Furthermore, as we have seen, there are reasons for selecting larger values for V_{OV}.
2. The numerator quantity V_A is both process- and device-dependent, and its value has been steadily decreasing.

As a result, the intrinsic gain realized in a single MOSFET amplifier stage fabricated in a modern short-channel technology is only 20 V/V to 40 V/V, at least an order of magnitude lower than that for a BJT.

The third formula given for A_0 in Table G.3 points out a very interesting fact: For a given process technology (V_A' and $\mu_n C_{ox}$) and a given device (L and W), the intrinsic gain is inversely proportional to $\sqrt{I_D}$. This is illustrated in Fig. G.1, which shows a typical plot of A_0 versus the bias current I_D. The plot confirms that the gain increases as the bias current is lowered. The gain, however, levels off at very low currents. This is because the MOSFET enters the subthreshold region of operation (Section 5.1.9), where it becomes very much like a BJT with an exponential current–voltage characteristic. The intrinsic gain then becomes constant, just like that of a BJT. Note, however, that although a higher gain is achieved at lower bias currents, the price paid is a lower g_m and less ability to drive capacitive loads and thus a decrease in bandwidth. This point will be further illustrated shortly.

Example G.2

We wish to compare the values of g_m, input resistance at the gate (base), r_o, and A_0 for an NMOS transistor fabricated in the 0.25-µm technology specified in Table G.1 and an *npn* transistor fabricated in the

Example G.2 *continued*

low-voltage technology specified in Table G.2. Assume both devices are operating at a drain (collector) current of 100 μA. For the MOSFET, let $L = 0.4$ μm and $W = 4$ μm, and specify the required V_{OV}.

Solution

For the NMOS transistor,

$$I_D = \frac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)V_{OV}^2$$

$$100 = \frac{1}{2} \times 267 \times \frac{4}{0.4} \times V_{OV}^2$$

Thus,

$$V_{OV} = 0.27 \text{ V}$$

$$g_m = \sqrt{2(\mu_n C_{ox})\left(\frac{W}{L}\right)I_D}$$

$$= \sqrt{2 \times 267 \times 10 \times 100} = 0.73 \text{ mA/V}$$

$$R_{\text{in}} = \infty$$

$$r_o = \frac{V_A' L}{I_D} = \frac{5 \times 0.4}{0.1} = 20 \text{ k}\Omega$$

$$A_0 = g_m r_o = 0.73 \times 20 = 14.6 \text{ V/V}$$

For the *npn* transistor,

$$g_m = \frac{I_C}{V_T} = \frac{0.1 \text{ mA}}{0.025 \text{ V}} = 4 \text{ mA/V}$$

$$R_{\text{in}} = r_\pi = \beta_0/g_m = \frac{100}{4 \text{ mA/V}} = 25 \text{ k}\Omega$$

$$r_o = \frac{V_A}{I_C} = \frac{35}{0.1 \text{ mA}} = 350 \text{ k}\Omega$$

$$A_0 = g_m r_o = 4 \times 350 = 1400 \text{ V/V}$$

EXERCISE

G.2 For an NMOS transistor fabricated in the 0.5-μm process specified in Table G.1 with $W = 5$ μm and $L = 0.5$ μm, find the transconductance and the intrinsic gain obtained at $I_D = 10$ μA, 100 μA, and 1 mA.
Ans. 0.2 mA/V, 200 V/V; 0.6 mA/V, 62 V/V; 2 mA/V, 20 V/V

G.3.7 High-Frequency Operation

The simplified high-frequency equivalent circuits for the MOSFET and the BJT are very similar, and so are the formulas for determining their unity-gain frequency (also called **transition frequency**) f_T. As we demonstrate in Chapter 10, f_T is a measure of the *intrinsic* bandwidth of the transistor itself and does *not* take into account the effects of capacitive loads. We address the issue of capacitive loads shortly. For the time being, note the striking similarity between the approximate formulas given in Table G.3 for the value of f_T of the two devices. In both cases f_T is inversely proportional to the square of the critical dimension of the device: the channel length for the MOSFET and the base width for the BJT. These formulas also clearly indicate that shorter-channel MOSFETs[4] and narrower-base BJTs are inherently capable of a wider bandwidth of operation. It is also important to note that while for the BJT the approximate expression for f_T indicates that it is entirely process determined, the corresponding expression for the MOSFET shows that f_T is proportional to the overdrive voltage V_{OV}. Thus we have conflicting requirements on V_{OV}: While a higher low-frequency gain is achieved by operating at a low V_{OV}, wider bandwidth requires an increase in V_{OV}. Therefore the selection of a value for V_{OV} involves, among other considerations, a trade-off between gain and bandwidth.

For *npn* transistors fabricated in the modern low-voltage process, f_T is in the range of 10 GHz to 20 GHz as compared to the 400 MHz to 600 MHz obtained with the standard high-voltage process. In the MOS case, NMOS transistors fabricated in a modern submicron technology, such as the 0.18-μm process, achieve f_T values in the range of 5 GHz to 15 GHz.

Before leaving the subject of high-frequency operation, let's look into the effect of a capacitive load on the bandwidth of the common-source (common-emitter) amplifier. For this purpose we shall assume that the frequencies of interest are much lower than f_T of the transistor. Hence we shall not take the transistor capacitances into account. Figure G.2(a) shows a common-source amplifier with a capacitive load C_L. The voltage gain from gate to drain can be found as follows:

$$V_o = -g_m V_{gs}(r_o \parallel C_L)$$

$$= -g_m V_{gs} \frac{r_o \dfrac{1}{sC_L}}{r_o + \dfrac{1}{sC_L}}$$

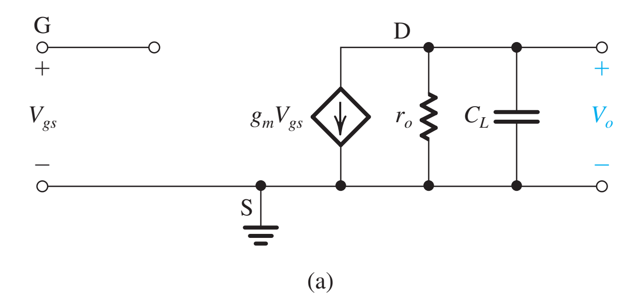
$$A_v = \frac{V_o}{V_{gs}} = -\frac{g_m r_o}{1 + sC_L r_o} \tag{G.1}$$

Thus the gain has, as expected, a low-frequency value of $g_m r_o = A_0$ and a frequency response of the single-time-constant (STC) low-pass type with a break (pole) frequency at

$$\omega_P = \frac{1}{C_L r_o} \tag{G.2}$$

Obviously this pole is formed by r_o and C_L. A sketch of the magnitude of gain versus frequency is shown in Fig. G.2(b). We observe that the gain crosses the 0-dB line at

[4]Although the reason is beyond our capabilities at this stage, f_T of MOSFETs that have very short channels varies inversely with L rather than with L^2.



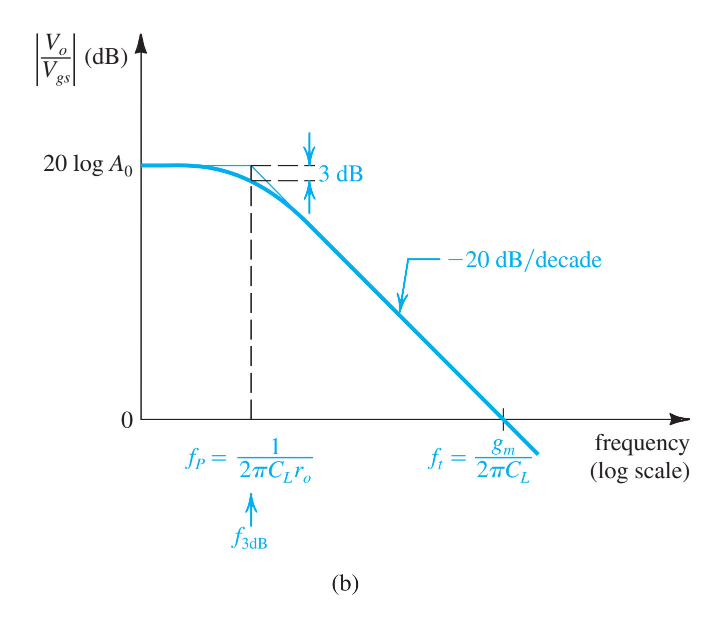


Figure G.2 Frequency response of a CS amplifier loaded with a capacitance C_L and fed with an ideal voltage source. It is assumed that the transistor is operating at frequencies much lower than f_T, and thus the internal capacitances are not taken into account.

frequency ω_t,

$$\omega_t = A_0\omega_P = (g_m r_o)\frac{1}{C_L r_o}$$

Thus,

$$\omega_t = \frac{g_m}{C_L} \tag{G.3}$$

That is, the **unity-gain frequency** or, equivalently, the **gain–bandwidth product**[5] ω_t is the ratio of g_m and C_L. We thus clearly see that for a given capacitive load C_L, a larger gain–bandwidth product is achieved by operating the MOSFET at a higher g_m. Identical analysis and conclusions apply to the case of the BJT. In each case, bandwidth increases as bias current is increased.

[5]The unity-gain frequency and the gain–bandwidth product of an amplifier are the same when the frequency response is of the single-pole type; otherwise the two parameters may differ.

G.3.8 Design Parameters

For the BJT there are three design parameters—I_C, V_{BE}, and I_S (or, equivalently, the area of the emitter–base junction)—and the designer can select any two. However, since I_C is exponentially related to V_{BE} and is very sensitive to the value of V_{BE} (V_{BE} changes by only 60 mV for a factor of 10 change in I_C), I_C is much more useful than V_{BE} as a design parameter. As mentioned earlier, the utility of the EBJ area as a design parameter is rather limited because of the narrow range over which A_E can vary. It follows that for the BJT there is only one effective design parameter: the collector current I_C. Finally, note that we have not considered V_{CE} to be a design parameter, since its effect on I_C is only secondary. Of course, as we learned in Chapter 7, V_{CE} affects the output-signal swing.

For the MOSFET there are four design parameters—I_D, V_{OV}, L, and W—and the designer can select any three. For analog circuit applications the trade-off in selecting a value for L is between the higher speed of operation (wider amplifier bandwidth) obtained at lower values of L and the higher intrinsic gain obtained at larger values of L. Usually one selects an L of about 25% to 50% greater than $L_{\min}$.

The second design parameter is V_{OV}. We have already made numerous remarks about the effect of the value of V_{OV} on performance. Usually, for submicron technologies, V_{OV} is selected in the range of 0.1 V to 0.3 V.

Once values for L and V_{OV} have been selected, the designer is left with the selection of the value of I_D or W (or, equivalently, W/L). For a given process and for the selected values of L and V_{OV}, I_D is proportional to W/L. It is important to note that the choice of I_D or, equivalently, of W/L has no bearing on the value of intrinsic gain A_0 and the transition frequency f_T. However, it affects the value of g_m and hence the gain–bandwidth product. Figure G.3 illustrates this point by showing how the gain of a common-source amplifier operated at a constant V_{OV} varies with I_D (or, equivalently, W/L). Note that while the dc gain remains unchanged, increasing W/L and, correspondingly, I_D, increases the bandwidth proportionally. This, however, assumes that the load capacitance C_L is not affected by the device size, an assumption that may not be entirely justified in some cases.

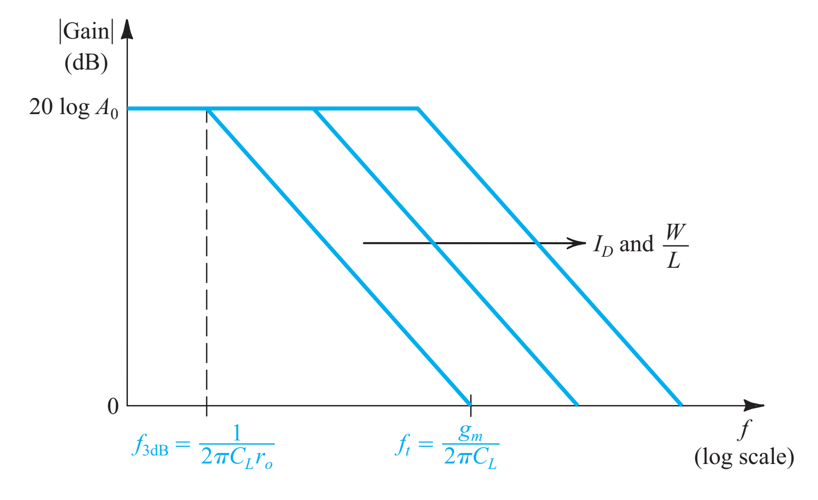


Figure G.3 Increasing I_D or W/L increases the bandwidth of a MOSFET amplifier operated at a constant V_{OV} and loaded by a constant capacitance C_L.

Example G.3

In this example we investigate the gain and the high-frequency response of an *npn* transistor and an NMOS transistor. For the *npn* transistor, assume that it is fabricated in the low-voltage process specified in Table G.2, and assume that $C_\mu \simeq C_{\mu 0}$. For $I_C = 10\ \mu A$, 100 µA, and 1 mA, find g_m, r_o, A_0, C_{de}, C_{je}, C_π, C_μ, and f_T. Also, for each value of I_C, find the gain–bandwidth product f_t of a common-emitter amplifier loaded by a 1-pF capacitance, neglecting the internal capacitances of the transistor. For the NMOS transistor, assume that it is fabricated in the 0.25-µm CMOS process with $L = 0.4\ \mu m$. Let the transistor be operated at $V_{OV} = 0.25$ V. Find W/L that is required to obtain $I_D = 10\ \mu A$, 100 µA, and 1 mA. At each value of I_D, find g_m, r_o, A_0, C_{gs}, C_{gd}, and f_T. Also, for each value of I_D, determine the gain–bandwidth product f_T of a common-source amplifier loaded by a 1-pF capacitance, neglecting the internal capacitances of the transistor.

Solution

For the *npn* transistor,

$$g_m = \frac{I_C}{V_T} = \frac{I_C}{0.025} = 40 I_C \text{ A/V}$$

$$r_o = \frac{V_A}{I_C} = \frac{35}{I_C}\ \Omega$$

$$A_0 = \frac{V_A}{V_T} = \frac{35}{0.025} = 1400 \text{ V/V}$$

$$C_{de} = \tau_F g_m = 10 \times 10^{-12} \times 40 I_C = 0.4 \times 10^{-9} I_C \text{ F}$$

$$C_{je} \simeq 2C_{je0} = 10 \text{ fF}$$

$$C_\pi = C_{de} + C_{je}$$

$$C_\mu \simeq C_{\mu 0} = 5 \text{ fF}$$

$$f_T = \frac{g_m}{2\pi (C_\pi + C_\mu)}$$

$$f_t = \frac{g_m}{2\pi C_L} = \frac{g_m}{2\pi \times 1 \times 10^{-12}}$$

We thus obtain the following results:

I_C	g_m (mA/V)	r_o (kΩ)	A_0 (V/V)	C_{de} (fF)	C_{je} (fF)	C_π (fF)	C_μ (fF)	f_T (GHz)	f_t (MHz)
10 µA	0.4	3500	1400	4	10	14	5	3.4	64
100 µA	4	350	1400	40	10	50	5	11.6	640
1 mA	40	35	1400	400	10	410	5	15.3	6400

Example G.3 *continued*

For the NMOS transistor,

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}V_{OV}^2$$
$$= \frac{1}{2} \times 267 \times \frac{W}{L} \times \frac{1}{16}$$

Thus,

$$\frac{W}{L} = 0.12 I_D$$
$$g_m = \frac{I_D}{V_{OV}/2} = \frac{I_D}{0.25/2} = 8 I_D \text{ A/V}$$
$$r_o = \frac{V_A' L}{I_D} = \frac{5 \times 0.4}{I_D} = \frac{2}{I_D}\ \Omega$$
$$A_0 = g_m r_o = 16 \text{ V/V}$$
$$C_{gs} = \frac{2}{3}WLC_{ox} + C_{ov} = \frac{2}{3}W \times 0.4 \times 5.8 + 0.6W$$
$$C_{gd} = C_{ov} = 0.6W$$
$$f_T = \frac{g_m}{2\pi\left(C_{gs} + C_{gd}\right)}$$
$$f_t = \frac{g_m}{2\pi C_L}$$

We thus obtain the following results:

I_D	W/L	g_m (mA/V)	r_o (kΩ)	A_0 (V/V)	C_{gs} (fF)	C_{gd} (fF)	f_T (GHz)	f_t (MHz)
10 μA	1.2	0.08	200	16	1.03	0.29	9.7	12.7
100 μA	12	0.8	20	16	10.3	2.9	9.7	127
1 mA	120	8	2	16	103	29	9.7	1270

EXERCISE

G.3 Find I_D, g_m, r_o, A_0, C_{gs}, C_{gd}, and f_T for an NMOS transistor fabricated in the 0.5-μm CMOS technology specified in Table G.1. Let $L = 0.5$ μm, $W = 5$ μm, and $V_{OV} = 0.3$ V.
Ans. 85.5 μA; 0.57 mA/V; 117 kΩ; 66.7 V/V; 8.3 fF; 2 fF; 8.8 GHz

G.4 Combining MOS and Bipolar Transistors—BiCMOS Circuits

From the discussion above it should be evident that the BJT has the advantage over the MOSFET of a much higher transconductance (g_m) at the same value of dc bias current. Thus, in addition to realizing higher voltage gains per amplifier stage, bipolar transistor amplifiers have superior high-frequency performance compared to their MOS counterparts.

On the other hand, the practically infinite input resistance at the gate of a MOSFET makes it possible to design amplifiers with extremely high input resistances and an almost zero input bias current. Also, as mentioned earlier, the MOSFET provides an excellent implementation of a switch, a fact that has made CMOS technology capable of realizing a host of analog circuit functions that are not possible with bipolar transistors.

It can thus be seen that each of the two transistor types has its own distinct and unique advantages: Bipolar technology has been extremely useful in the design of very-high-quality general-purpose circuit building blocks, such as op amps. On the other hand, CMOS, with its very high packing density and its suitability for both digital and analog circuits, has become the technology of choice for the implementation of very-large-scale integrated circuits. Nevertheless, the performance of CMOS circuits can be improved if the designer has available (on the same chip) bipolar transistors that can be employed in functions that require their high g_m and excellent current-driving capability. A technology that allows the fabrication of high-quality bipolar transistors on the same chip as CMOS circuits is aptly called **BiCMOS**. At appropriate locations throughout this book we present interesting and useful BiCMOS circuit blocks.

G.5 Validity of the Square-Law MOSFET Model

We conclude this appendix with a comment on the validity of the simple square-law model we have been using to describe the operation of the MOS transistor. While this simple model works well for devices with relatively long channels ($>1\ \mu$m), it does *not* provide an accurate representation of the operation of short-channel devices. This is because a number of physical phenomena come into play in these submicron devices, resulting in what are called **short-channel effects**. Although a detailed study of short-channel effects is beyond the scope of this book, it should be mentioned that MOSFET models have been developed that take these effects into account. However, they are understandably quite complex and do not lend themselves to hand analysis of the type needed to develop insight into circuit operation. Rather, these models are suitable for computer simulation and are indeed used in SPICE (Appendix B). For quick, manual analysis, however, we will continue to use the square-law model, which is the basis for the comparison of Table G.3.

G.1 Find the range of I_D obtained in a particular NMOS transistor as its overdrive voltage is increased from 0.15 V to 0.4 V. If the same range is required in I_C of a BJT, what is the corresponding change in V_{BE}?

G.2 What range of I_C is obtained in an *npn* transistor as a result of changing the area of the emitter–base junction by a factor of 10 while keeping V_{BE} constant? If I_C is to be kept constant, by what amount must V_{BE} change?

G.3 For each of the CMOS technologies specified in Table G.1, find the $|V_{OV}|$ and hence the $|V_{GS}|$ required to operate a device with a W/L of 10 at a drain current $I_D =$ 100 μA. Ignore channel-length modulation.

G.4 Consider NMOS and PMOS devices fabricated in the 0.25-μm process specified in Table G.1. If both devices are to operate at $|V_{OV}| = 0.25$ V and $I_D = 100$ μA, what must their W/L ratios be?

G.5 Consider NMOS and PMOS transistors fabricated in the 0.25-μm process specified in Table G.1. If the two devices are to be operated at equal drain currents, what must the ratio of $(W/L)_p$ to $(W/L)_n$ be to achieve equal values of g_m?

G.6 An NMOS transistor fabricated in the 0.18-μm CMOS process specified in Table G.1 is operated at $V_{OV} = 0.2$ V. Find the required W/L and I_D to obtain a g_m of 10 mA/V. At what value of I_C must an *npn* transistor be operated to achieve this value of g_m?

G.7 For each of the CMOS process technologies specified in Table G.1, find the g_m of an NMOS and a PMOS transistor with $W/L = 10$ operated at $I_D = 100$ μA.

G.8 An NMOS transistor operated with an overdrive voltage of 0.25 V is required to have a g_m equal to that of an *npn* transistor operated at $I_C = 0.1$ mA. What must I_D be? What value of g_m is realized?

G.9 It is required to find the incremental (i.e., small-signal) resistance of each of the diode-connected transistors shown in Fig. PG.9. Assume that the dc bias current $I = 0.1$ mA. For the MOSFET, let $\mu_n C_{ox} = 200$ μA/V^2 and $W/L = 10$.

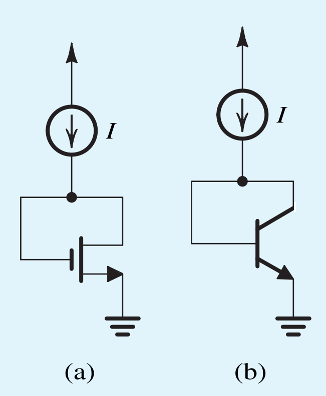


Figure PG.9

G.10 For an NMOS transistor with $L = 1$ μm fabricated in the 0.8-μm process specified in Table G.1, find g_m, r_o, and A_0 if the device is operated with $V_{OV} = 0.5$ V and $I_D = 100$ μA. Also, find the required device width W.

G.11 For an NMOS transistor with $L = 0.3$ μm fabricated in the 0.18-μm process specified in Table G.1, find g_m, r_o, and A_0 obtained when the device is operated at $I_D = 100$ μA with $V_{OV} = 0.2$ V. Also, find W.

G.12 Fill in the table below. For the BJT, let $\beta = 100$ and $V_A = 100$ V. For the MOSFET, let $\mu_n C_{ox} = 200$ μA/V^2, $W/L = 40$, and $V_A = 10$ V. Note that R_{in} refers to the input resistance at the control input terminal (gate, base) with the (source, emitter) grounded.

	BJT		MOSFET	
Bias Current	$I_C = 0.1$ mA	$I_C = 1$ mA	$I_D = 0.1$ mA	$I_D = 1$ mA
g_m (mA/V)				
r_o (kΩ)				
A_0 (V/V)				
R_{in} (kΩ)				

G.13 For an NMOS transistor fabricated in the 0.18-μm process specified in Table G.1 with $L = 0.3$ μm and $W =$ 6 μm, find the value of f_T obtained when the transistor is operated at $V_{OV} = 0.2$ V. Use both the formula in terms of C_{gs} and C_{gd} and the approximate formula. Why does the approximate formula overestimate f_T?

G.14 An NMOS transistor fabricated in the 0.18-μm process specified in Table G.1 and having $L = 0.3$ μm and $W = 6$ μm is operated at $V_{OV} = 0.2$ V and used to drive a

capacitive load of 100 fF. Find A_0, f_P (or $f_{3\text{ dB}}$), and f_t. At what I_D value is the transistor operating? If it is required to double f_t, what must I_D become? What happens to A_0 and f_P in this case?

G.15 For an *npn* transistor fabricated in the high-voltage process specified in Table G.2, evaluate f_T at $I_C = 10\ \mu\text{A}$, $100\ \mu\text{A}$, and 1 mA. Assume $C_\mu \simeq C_{\mu 0}$. Repeat for the low-voltage process.

G.16 Consider an NMOS transistor fabricated in the 0.8-μm process specified in Table G.1. Let the transistor have $L = 1\ \mu\text{m}$, and assume it is operated at $I_D = 100\ \mu\text{A}$.

(a) For $V_{OV} = 0.25$ V, find W, g_m, r_o, A_0, C_{gs}, C_{gd}, and f_T.
(b) To what must V_{OV} be changed to double f_T? Find the new values of W, g_m, r_o, A_0, C_{gs}, and C_{gd}.

G.17 For a lateral *pnp* transistor fabricated in the high-voltage process specified in Table G.2, find f_T if the device is operated at a collector bias current of 1 mA. Compare to the value obtained for a vertical *npn*.

G.18 Show that for a MOSFET the selection of L and V_{OV} determines A_0 and f_T. In other words, show that A_0 and f_T will not depend on I_D and W.

G.19 Consider an NMOS transistor fabricated in the 0.18-μm technology specified in Table G.1. Let the transistor be operated at $V_{OV} = 0.2$ V. Find A_0 and f_T for $L = 0.2\ \mu\text{m}$, 0.3 μm, and 0.4 μm.

D G.20 Consider an NMOS transistor fabricated in the 0.5-μm process specified in Table G.1. Let $L = 0.5\ \mu\text{m}$ and $V_{OV} = 0.3$ V. If the MOSFET is connected as a common-source amplifier with a load capacitance $C_L = 1$ pF (as in Fig. G.2a), find the required transistor width W and bias current I_D to obtain a unity-gain bandwidth of 100 MHz. Also, find A_0 and $f_{3\text{ dB}}$.

APPENDIX H

DESIGN OF STAGGER-TUNED AMPLIFIERS

As mentioned in Section 17.11.7, a much better overall response of a tuned amplifier is obtained by stagger-tuning (as opposed to synchronous tuning) the individual stages, as illustrated in Fig. 17.46 which is repeated here as Fig. H.1. Stagger-tuned amplifiers are usually designed so that the overall response exhibits *maximal flatness* around the center-frequency f_0. Such a response can be obtained by transforming the response of a maximally flat (Butterworth) low-pass filter up the frequency axis to ω_0. We show here how this can be done.

The transfer function of a second-order bandpass filter can be expressed in terms of its poles as

$$T(s) = \frac{a_1 s}{\left(s + \frac{\omega_0}{2Q} - j\omega_0\sqrt{1 - \frac{1}{4Q^2}}\right)\left(s + \frac{\omega_0}{2Q} + j\omega_0\sqrt{1 - \frac{1}{4Q^2}}\right)} \tag{H.1}$$

For a narrow-band filter, $Q \gg 1$, and for values of s in the neighborhood of $+j\omega_0$ (see Fig. H.2b), the second factor in the denominator is approximately $(s + j\omega_0 \simeq 2s)$. Hence Eq. (H.1) can be approximated in the neighborhood of $j\omega_0$ by

$$T(s) \simeq \frac{a_1/2}{s + \omega_0/2Q - j\omega_0} = \frac{a_1/2}{(s - j\omega_0) + \omega_0/2Q} \tag{H.2}$$

This is known as the **narrow-band approximation**.[1] Note that the magnitude response, for $s = j\omega$, has a peak value of $a_1 Q/\omega_0$ at $\omega = \omega_0$, as expected.

Now consider a first-order low-pass network with a single pole at $p = -\omega_0/2Q$ (we use p to denote the complex frequency variable for the low-pass filter). Its transfer function is

$$T(p) = \frac{K}{p + \omega_0/2Q} \tag{H.3}$$

where K is a constant. Comparing Eqs. (H.2) and (H.3) we note that they are identical for $p = s - j\omega_0$ or, equivalently,

$$s = p + j\omega_0 \tag{H.4}$$

[1]The bandpass response is *geometrically symmetrical* around the center frequency ω_0. That is, each pair of frequencies ω_1 and ω_2 at which the magnitude response is equal are related by $\omega_1\omega_2 = \omega_0^2$. For high Q, the symmetry becomes almost *arithmetic* for frequencies close to ω_0. That is, two frequencies with the same magnitude response are almost equally spaced from ω_0. The same is true for higher-order bandpass filters designed using the transformation presented in this section.

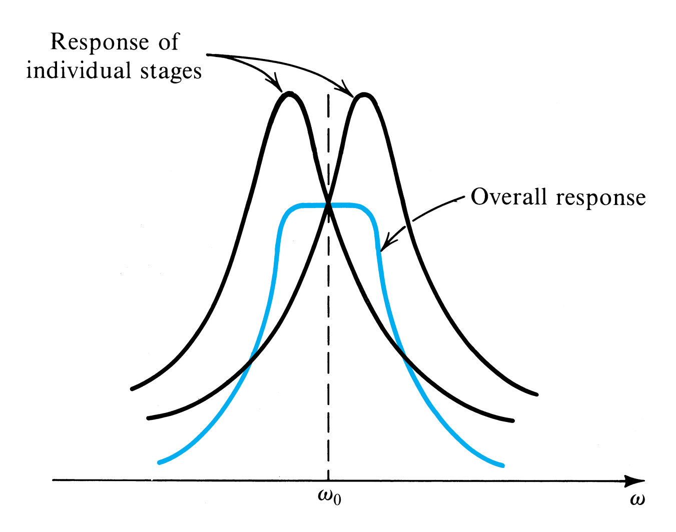


Figure H.1 Stagger-tuning the individual resonant circuits can result in an overall response with a passband flatter than that obtained with synchronous tuning (Fig. 17.48).

This result implies that the response of the second-order bandpass filter *in the neighborhood of its center frequency* $s = j\omega_0$ is identical to the response of a first-order low-pass filter with a pole at $(-\omega_0/2Q)$ *in the neighborhood of* $p = 0$. Thus the bandpass response can be obtained by shifting the pole of the low-pass prototype and adding the complex-conjugate pole, as illustrated in Fig. H.2(b). This is called a **lowpass-to-bandpass transformation** for *narrow-band* filters.

The transformation $p = s - j\omega_0$ can be applied to low-pass filters of order greater than one. For instance, we can transform a maximally flat, second-order low-pass filter $(Q = 1/\sqrt{2})$ to obtain a maximally flat bandpass filter. If the 3-dB bandwidth of the bandpass filter is to be B rad/s, then the low-pass filter should have a 3-dB frequency (and thus a pole frequency) of $(B/2)$ rad/s, as illustrated in Fig. H.3. The resulting fourth-order bandpass filter will be a stagger-tuned one, with its two tuned circuits (refer to Fig. 16.48) having

$$\omega_{01} = \omega_0 + \frac{B}{2\sqrt{2}} \qquad B_1 = \frac{B}{\sqrt{2}} \qquad Q_1 \simeq \frac{\sqrt{2}\omega_0}{B} \tag{H.5}$$

$$\omega_{02} = \omega_0 - \frac{B}{2\sqrt{2}} \qquad B_2 = \frac{B}{\sqrt{2}} \qquad Q_2 = \frac{\sqrt{2}\omega_0}{B} \tag{H.6}$$

Note that for the overall response to have a normalized center-frequency gain of unity, the individual responses have to have equal center-frequency gains of $\sqrt{2}$, as shown in Fig. H.3(d).

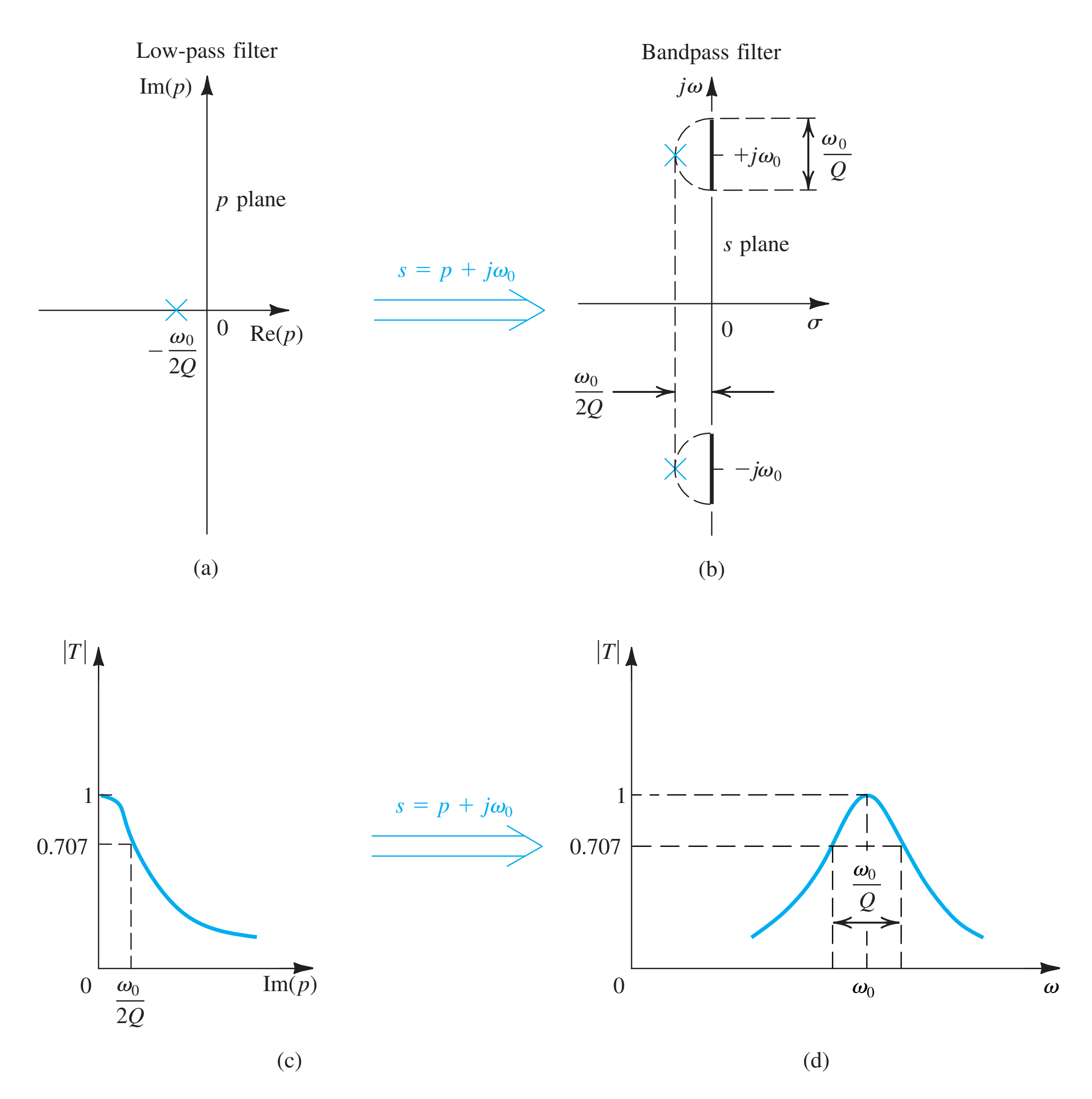


Figure H.2 Obtaining a second-order narrow-band bandpass filter by transforming a first-order low-pass filter. **(a)** Pole of the first-order filter in the *p* plane. **(b)** Applying the transformation $s = p + j\omega_0$ and adding a complex-conjugate pole results in the poles of the second-order bandpass filter. **(c)** Magnitude response of the first-order low-pass filter. **(d)** Magnitude response of the second-order bandpass filter.

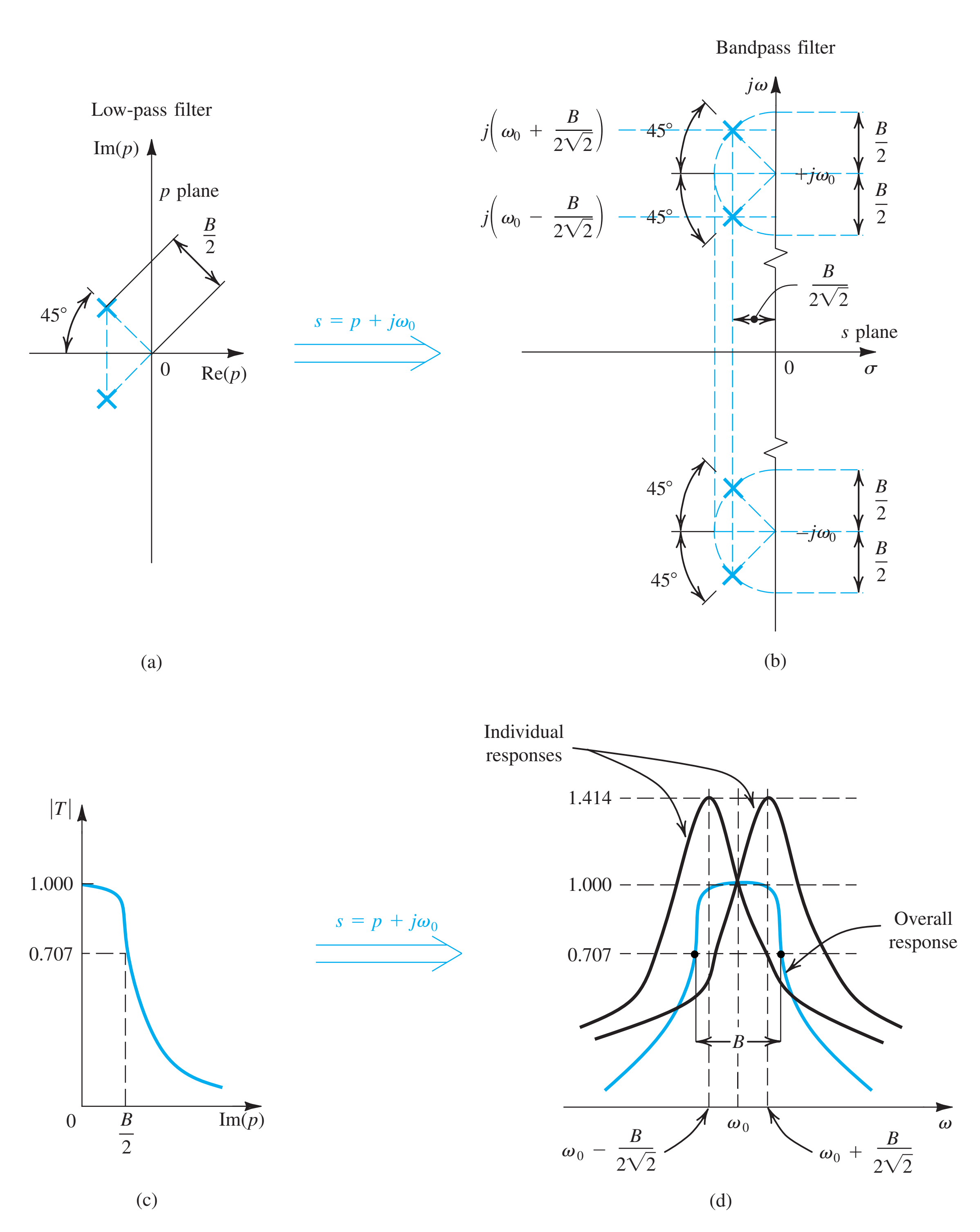


Figure H.3 Obtaining the poles and the frequency response of a fourth-order stagger-tuned, narrow-band bandpass amplifier by transforming a second-order low-pass, maximally flat response.

EXERCISES

D H.1 A stagger-tuned design for the IF amplifier specified in Exercise 17.36 is required. Find f_{01}, B_1, f_{02}, and B_2. Also give the value of C and R for each of the two stages. (Recall that 3-μH inductors are to be used.)

Ans. 10.77 MHz; 141.4 kHz; 10.63 MHz; 141.4 kHz; 72.8 pF; 15.5 kΩ; 74.7 pF; 15.1 kΩ

H.2 Using the fact that the voltage gain at resonance is proportional to the value of R, find the ratio of the gain at 10.7 MHz of the stagger-tuned amplifier designed in Exercise H.1 and the synchronously tuned amplifier designed in Exercise 17.36. (*Hint:* For the stagger-tuned amplifier, note that the gain at ω_0 is equal to the product of the gains of the individual stages at their 3-dB frequencies.)

Ans. 2.42

PROBLEMS

***H.1** This problem investigates the selectivity of maximally flat stagger-tuned amplifiers derived in the manner illustrated in Fig. H.3.

(a) The low-pass maximally flat (Butterworth) filter having a 3-dB bandwidth $B/2$ and order N has the magnitude response

$$|T| = 1\Bigg/\sqrt{1+\left(\frac{\Omega}{B/2}\right)^{2N}}$$

where $\Omega = \text{Im}(p)$ is the frequency in the low-pass domain. (This relationship can be obtained using the information provided in Section 17.3 on Butterworth filters.) Use this expression to obtain for the corresponding bandpass filter at $\omega = \omega_0 + \delta\omega$, where $\delta\omega \ll \omega_0$, the relationship

$$|T| = 1\Bigg/\sqrt{1+\left(\frac{\delta\omega}{B/2}\right)^{2N}}$$

(b) Use the transfer function in (a) to find the attenuation (in decibels) obtained at a bandwidth of $2B$ for $N = 1$ to 5. Also find the ratio of the 30-dB bandwidth to the 3-dB bandwidth for $N = 1$ to 5.

****H.2** Consider a sixth-order, stagger-tuned bandpass amplifier with center frequency ω_0 and 3-dB bandwidth B. The poles are to be obtained by shifting those of the third-order maximally flat low-pass filter, given in Fig. 17.10(c). For each of the three resonant circuits, find ω_0, the 3-dB bandwidth, and Q.

APPENDIX I

BIBLIOGRAPHY

HISTORY OF ELECTRONICS

L. Berlin, *The Man Behind the Microchip: Robert Noyce and the Invention of Silicon Valley*, New York: Oxford University Press, 2005.

J. Gertner, *The Idea Factory: Bell Labs and the Great Age of American Innovation*, New York: The Penguin Press, 2012.

IEEE Solid-State Circuits Magazine, and its predecessor, *IEEE Solid-State Circuits Newsletter*, published quarterly by the IEEE Solid-State Circuits Society.

T.R. Reid, *The Chip*, New York: Random House, 2001.

J.N. Shurkin, *Broken Genius: The Rise and Fall of William Shockley, Creator of the Electronic Age*, New York, Macmillan, 2008.

J. Williams, editor, *Analog Circuit Design: Art, Science, and Personalities*, Boston: Butterworth-Heinemann, 1991.

GENERAL TEXTBOOKS ON ELECTRONIC CIRCUITS

E.F. Angelo Jr., *Electronics: BJTs, FETs, and Microcircuits*, New York: McGraw-Hill, 1969.

S.B. Burns and P.R. Bond, *Principles of Electronic Cir*cuits, St. Paul: West, 1987.

M.S. Ghausi, *Electronic Devices and Circuits: Discrete and Integrated,* New York: Holt, Rinehart and Winston, 1985.

P.E. Gray and C.L. Searle, *Electronic Principles,* New York: Wiley, 1969.

A.R. Hambley, *Electronics, 2nd ed.,* Upper Saddle River, NJ: Prentice-Hall, 1999.

W.H. Hayt and G.W. Neudeck, *Electronic Circuit Analysis and Design*, *2nd ed.,* Boston: Houghton Mifflin Co., 1984.

C.A. Holt, *Electronic Circuits,* New York: Wiley, 1978.

M.N. Horenstein, *Microelectronic Circuits and Devices*, *2nd ed.*, Englewood Cliffs, NJ: Prentice-Hall, 1995.

R.T. Howe and C.G. Sodini, *Microelectronics—An Integrated Approach*, Englewood Cliffs, NJ: Prentice-Hall, 1997.

R.C. Jaeger and T.N. Blalock, *Microelectronic Circuit Design, 4th ed.,* New York: McGraw-Hill, 2011.

N.R. Malik, *Electronic Circuits: Analysis, Simulation, and Design*, Englewood Cliffs, NJ: Prentice-Hall, 1995.

J. Millman and A. Grabel, *Microelectronics, 2nd ed.,* New York: McGraw-Hill, 1987.

D.A. Neamen, *Electronic Circuit Analysis and Design, 4th ed.,* New York: McGraw-Hill, 2010.

M.H. Rashid, *Microelectronic Circuits: Analysis and Design,* Boston: PWS, 1999.

B. Ravazi, *Fundamentals of Microelectronics*, 2nd ed. Hoboken, NJ: Wiley, 2014.

D.L. Schilling and C. Belove, *Electronic Circuits, 2nd ed.,* New York: McGraw-Hill, 1979.

R.A. Spencer and M.S. Ghausi, *Introduction to Electronic Circuit Design,* Upper Saddle River, NJ: Pearson Education Inc. (Prentice-Hall), 2003.

CIRCUIT AND SYSTEM ANALYSIS

L.S. Bobrow, *Elementary Linear Circuit Analysis, 2nd ed.,* New York: Holt, Rinehart and Winston, 1987.

A.M. Davis, *Linear Circuit Analysis,* Boston, PWS Publishing Company, 1998.

S.S. Haykin, *Active Network Theory,* Reading, MA: Addison-Wesley, 1970.

W.H. Hayt, G.E. Kemmerly, and S.M. Durbin, *Engineering Circuit Analysis*, *6th ed.*, New York: McGraw-Hill, 2003.

D. Irwin, *Basic Engineering Circuit Analysis, 7th ed.,* New York: Wiley, 2001.

B.P. Lathi, *Linear Systems and Signals,* New York: Oxford University Press, 1992.

J.W. Nilsson and S. Riedel, *Electronic Circuits, 7th ed.,* Upper Saddle River, NJ: Prentice-Hall, 2005.

DEVICES AND IC FABRICATION

R.S.C. Cobbold, *Theory and Applications of Field Effect Transistors*, New York: Wiley, 1969.

I. Getreu, *Modeling the Bipolar Transistor,* Beaverton, OR: Tektronix, Inc., 1976.

R.S. Muller and T.I. Kamins, *Device Electronics for Integrated Circuits*, *3rd ed.,* New York: Wiley, 2003.

J.D. Plummer, M.D. Deal, and P.B. Griffin, *Silicon VLSI Technology,* Upper Saddle River, NJ: Prentice-Hall, 2000.

D.L. Pulfrey and N.G. Tarr, *Introduction to Microelectronic Devices*, Englewood Cliffs, NJ: Prentice-Hall, 1989.

C.L. Searle, A.R. Boothroyd, E.J. Angelo, Jr., P.E. Gray, and D.O. Pederson, *Elementary Circuit Properties of Transistors*, Vol. 3 of the SEEC Series, New York: Wiley, 1964.

B.G. Streetman and S. Banerjee, *Solid-State Electronic Devices, 5th ed.,* Upper Saddle River, NJ: Prentice-Hall, 2000.

Y. Tsividis, *Operation and Modeling of the MOS Transistor*, *3rd ed.*, New York: Oxford University Press, 2011.

OPERATIONAL AMPLIFIERS

G.B. Clayton, *Experimenting with Operational Amplifiers*, London: Macmillan, 1975.

G.B. Clayton, *Operational Amplifiers, 2nd ed.,* London: Newnes-Butterworths, 1979.

S. Franco, *Design with Operational Amplifiers and Analog Integrated Circuits*, *3rd ed.*, New York: McGraw-Hill, 2001.

J.G. Graeme, G.E. Tobey, and L.P. Huelsman, *Operational Amplifiers: Design and Applications,* New York: McGraw-Hill, 1971.

W. Jung, *IC Op Amp Cookbook,* Indianapolis: Howard Sams, 1974.

E.J. Kennedy, *Operational Amplifier Circuits: Theory and Applications,* New York: Holt, Rinehart and Winston, 1988.

J.K. Roberge, *Operational Amplifiers: Theory and Practice,* New York: Wiley, 1975.

J.L. Smith, *Modern Operational Circuit Design,* New York: Wiley-Interscience, 1971.

J.V. Wait, L.P. Huelsman, and G.A. Korn, *Introduction to Operational Amplifiers*:*Theory and Applications*, New York: McGraw-Hill, 1975.

ANALOG CIRCUITS

P.E. Allen and D.R. Holberg, *CMOS Analog Circuit Design, 3rd ed.,* New York: Oxford University Press, 2012.

R. Jacob Baker, *CMOS: Circuit Design, Layout and Simulation, 3rd ed*. Hoboken, N.J.: Wiley, 2010.

K. Bult, *Transistor-Level Analog IC Design*. Notes for a short course organized by Mead, Ecole Polytechnique Féderal de Lausanne, 2002.

T. Chan Carusone, D.A. Johns, and K. Martin, *Analog Integrated Circuit Design, 2nd ed*. New York: Wiley, 2012.

M.J. Fonderic and J.H. Huising, *Design of Low-Voltage Bipolar Operational Amplifiers*, Boston: Kluwer Academic Publishers, 1993.

R.L. Geiger, P.E. Allen, and N.R. Strader, *VLSI Design Techniques for Analog and Digital Circuits,* New York: McGraw-Hill, 1990.

P.R. Gray, P.J. Hurst, S.H. Lewis, and R.G. Meyer, *Analysis and Design of Analog Integrated Circuits*, *5th ed.*, Hoboken, NJ: Wiley, 2008.

A.B. Grebene, *Bipolar and MOS Analog Integrated Circuit Design,* New York: Wiley, 1984.

R. Gregorian and G.C. Temes, *Analog MOS Integrated Circuits for Signal Processing,* New York: Wiley, 1986.

J.H. Huising, *Operational Amplifiers*, Boston: Kluwer Academic Publishers, 2001.

IEEE Journal of Solid-State Circuits, a monthly publication of the IEEE.

K. Laker and W. Sansen, *Design for Analog Integrated Circuits and Systems,* New York: McGraw-Hill, 1999.

H.S. Lee, "Analog Design," Chapter 8 in *BiCMOS Technology and Applications,* A.R. Alvarez, editor, Boston: Kluwer Academic Publishers, 1989.

National Semiconductor Corporation, Audio/Radio Handbook, Santa Clara, CA: National Semiconductor Corporation, 1980.

B. Razavi, *Design of Analog CMOS Integrated Cir*cuits, New York: McGraw-Hill, 2001.

J.K. Roberge, *Operational Amplifiers: Theory and Practice,* New York: Wiley, 1975.

S. Rosenstark, *Feedback Amplifier Principles,* New York: Macmillan, 1986.

W.M.C. Sansen, *Analog Design Essentials*, Dordrecht, The Netherlands: Springer, 2006.

A.S. Sedra and G.W. Roberts, "Current Conveyor Theory and Practice," Chapter 3 in *Analogue IC Design: The Current-Mode Approach*, C. Toumazou, F.J. Lidgey, and D.G. Haigh, editors, London: Peter Peregrinus, 1990.

R. Severns, editor, *MOSPOWER Applications Handbook*, Santa Clara, CA: Siliconix, 1984.

Texas Instruments, Inc., *Power Transistor and TTL Integrated-Circuit Applications*, New York: McGraw-Hill, 1977.

S. Soclof, *Applications of Analog Integrated Circuits,* Englewood Cliffs, NJ: Prentice-Hall, 1985.

J.M. Steininger, "Understanding wideband MOS transistors," *IEEE Circuits and Devices,* Vol. 6, No. 3, pp. 26–31, May 1990.

DIGITAL CIRCUITS

A.R. Alvarez, editor, *BiCMOS Technology and Applications*, *2nd ed.*, Boston: Kluwer Academic Publishers, 1993.

M.I. Elmasry, editor, *Digital MOS Integrated Circuits,* New York: IEEE Press, 1981. Also, *Digital MOS Integrated Circuits II*, 1992.

S.H.K Embabi, A. Bellaour, and M.I. Elmasry, *Digital BiCMOS Integrated Circuit Design,* Boston: Kluwer, 1993.

D.A. Hodges, H.G. Jackson, and R.A. Saleh, *Analysis and Design of Digital Integrated Circuits, 3rd ed.,* New York: McGraw-Hill, 2004.

IEEE Journal of Solid-State Circuits, a monthly publication of the IEEE.

S.M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits, 3rd ed.,* New York: McGraw-Hill, 2003.

R. Littauer, *Pulse Electronics,* New York: McGraw-Hill, 1965.

K. Martin, *Digital Integrated Circuit Design,* New York: Oxford University Press, 2000.

J. Millman and H. Taub, *Pulse, Digital, and Switching Waveforms,* New York: McGraw-Hill, 1965.

Motorola, *MECL Device Data,* Phoenix, AZ: Motorola Semiconductor Products, Inc., 1989.

Motorola, *MECL System Design Handbook,* Phoenix, AZ: Motorola Semiconductor Products, Inc., 1988.

J.M. Rabaey, *Digital Integrated Circuits,* Englewood Cliffs, NJ: Prentice-Hall, 1996. Note: A 2nd ed., with A. Chandrakasan and B. Nikolic, also appeared in 2003.

L. Strauss, *Wave Generation and Shaping, 2nd ed.,* New York: McGraw-Hill, 1970.

H. Taub and D. Schilling, *Digital Integrated Electronics*, New York: McGraw-Hill, 1977.

N. Weste and K. Eshraghian, *Principles of CMOS VLSI Design,* Reading, MA: Addison-Wesley, 1985 and 1993.

FILTERS AND TUNED AMPLIFIERS

P.E. Allen and E. Sanchez-Sinencio, *Switched-Capacitor Circuits*, New York: Van Nostrand Reinhold, 1984.

K.K. Clarke and D.T. Hess, *Communication Circuits: Analysis and Design,* Ch. 6, Reading, MA: Addison Wesley, 1971.

G. Daryanani, *Principles of Active Network Synthesis and Design,* New York: Wiley, 1976.

R. Gregorian and G.C. Temes, *Analog MOS Integrated Circuits for Signal Processing,* New York: Wiley-Interscience, 1986.

S.K. Mitra and C.F. Kurth, editors, *Miniaturized and Integrated Filters*, New York: Wiley-Interscience, 1989.

C. Ouslis and A. Sedra, "Designing custom filters," *IEEE Circuits and Devices,* May 1995, pp. 29–37.

R. Schaumann, M.S. Ghausi, and K.R. Laker, *Design of Analog Filters,* Englewood Cliffs, NJ: Prentice-Hall, 1990.

R. Schaumann, M. Soderstand, and K. Laker, editors, *Modern Active Filter Design,* New York: IEEE Press, 1981.

R. Schaumann, H. Xiao, and M.E. Van Valkenburg, *Design of Analog Filters, 2nd ed*. New York: Oxford University Press, 2010.

A.S. Sedra, "Switched-capacitor filter synthesis," in *MOS VLSI Circuits for Telecommunications,* Y. Tsividis and P. Antognetti, editors, Englewood Cliffs, NJ: Prentice-Hall, 1985.

A.S. Sedra and P.O. Brackett, *Filter Theory and Design: Active and Passive*, Portland, OR: Matrix, 1978.

M.E. Van Valkenburg, *Analog Filter Design,* New York: Holt, Rinehart and Winston, 1981.

A.I. Zverev, *Handbook of Filter Synthesis,* New York: Wiley, 1967.

SPICE

M.E. Herniter, *Schematic Capture with Cadence PSpice, 2nd ed.*, Upper Saddle River, NJ: Prentice-Hall, 2003.

G. Massobrio and P. Antognetti, *Semiconductor Device Modeling with SPICE, 2nd ed.,* New York: McGraw-Hill, 1993.

G.W. Roberts and A.S. Sedra, *SPICE*, New York: Oxford University Press, 1992 and 1997.

J.A. Svoboda, *PSpice for Linear Circuits*, New York: Wiley, 2002.

P.W. Tuinenga, SPICE: *A Guide to Circuit Simulation Analysis Using PSpice, 2nd ed.*, Englewood Cliffs, NJ: Prentice-Hall, 1992.

APPENDIX J

STANDARD RESISTANCE VALUES AND UNIT PREFIXES

Discrete resistors are available only in standard values. Table J.1 provides the multipliers for the standard values of 5%-tolerance and 1%-tolerance resistors. Thus, in the kilohm range of 5% resistors, one finds resistances of 1.0, 1.1, 1.2, 1.3, 1.5, ... In the same range, one finds 1% resistors of kilohm values of 1.00, 1.02, 1.05, 1.07, 1.10, ...

Table J.1 Standard Resistance Values

	1% Resistor Values (kΩ)			
5% Resistor Values (kΩ)	100–174	178–309	316–549	562–976
10	100	178	316	562
11	102	182	324	576
12	105	187	332	590
13	107	191	340	604
15	110	196	348	619
16	113	200	357	634
18	115	205	365	649
20	118	210	374	665
22	121	215	383	681
24	124	221	392	698
27	127	226	402	715
30	130	232	412	732
33	133	237	422	750
36	137	243	432	768
39	140	249	442	787
43	143	255	453	806
47	147	261	464	825
51	150	267	475	845
56	154	274	487	866
62	158	280	499	887
68	162	287	511	909
75	165	294	523	931
82	169	301	536	953
91	174	309	549	976

Table J.2 provides the SI unit prefixes used in this book and in all modern works in English.

Table J.2 SI Unit Prefixes

Name	Symbol	Factor
femto	f	$\times 10^{-15}$
pico	p	$\times 10^{-12}$
nano	n	$\times 10^{-9}$
micro	μ	$\times 10^{-6}$
milli	m	$\times 10^{-3}$
kilo	k	$\times 10^{3}$
mega	M	$\times 10^{6}$
giga	G	$\times 10^{9}$
tera	T	$\times 10^{12}$
peta	P	$\times 10^{15}$

Table J.3 provides the meter conversion factors.

Table J.3 Meter Conversion Factors

$1\ \mu\text{m} = 10^{-4}\ \text{cm} = 10^{-6}\ \text{m}$
$1\ \text{m} = 10^{2}\ \text{cm} = 10^{6}\ \mu\text{m} = 10^{9}\ \text{nm}$
$0.1\ \mu\text{m} = 100\ \text{nm}$
$1\ \text{Å} = 10^{-8}\ \text{cm} = 10^{-10}\ \text{m}$

APPENDIX K

TYPICAL PARAMETER VALUES FOR IC DEVICES FABRICATED IN CMOS AND BIPOLAR PROCESSES

Table K.1 Typical Values of CMOS Device Parameters

	0.8 μm		0.5 μm		0.25 μm		0.18 μm		0.13 μm		65 nm	
Parameter	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS
t_{ox} (nm)	15	15	9	9	6	6	4	4	2.7	2.7	1.4	1.4
C_{ox} (fF/μm^2)	2.3	2.3	3.8	3.8	5.8	5.8	8.6	8.6	12.8	12.8	25	25
μ (cm^2/V·s)	550	250	500	180	460	160	450	100	400	100	216	40
μC_{ox} (μA/V^2)	127	58	190	68	267	93	387	86	511	128	540	100
V_{t0} (V)	0.7	−0.7	0.7	−0.8	0.5	−0.6	0.5	−0.5	0.4	−0.4	0.35	−0.35
V_{DD} (V)	5	5	3.3	3.3	2.5	2.5	1.8	1.8	1.3	1.3	1.0	1.0
$\|V'_A\|$ (V/μm)	25	20	20	10	5	6	5	6	5	6	3	3
C_{ov} (fF/μm)	0.2	0.2	0.4	0.4	0.3	0.3	0.37	0.33	0.36	0.33	0.33	0.31

Table K.2 Typical Parameter Values for BJTs*

	Standard High-Voltage Process		Advanced Low-Voltage Process	
Parameter	*npn*	Lateral *pnp*	*npn*	Lateral *pnp*
A_E (μm^2)	500	900	2	2
I_S (A)	5×10^{-15}	2×10^{-15}	6×10^{-18}	6×10^{-18}
β_0 (A/A)	200	50	100	50
V_A (V)	130	50	35	30
V_{CEO} (V)	50	60	8	18
τ_F	0.35 ns	30 ns	10 ps	650 ps
C_{je0}	1 pF	0.3 PF	5 fF	14 fF
$C_{\mu 0}$	0.3 pF	1 pF	5 fF	15 fF
r_x (Ω)	200	300	400	200

*Adapted from Gray et al. (2001); see Appendix I.

Note: For more information, refer to Appendix G (on the book’s Website).

APPENDIX L

ANSWERS TO SELECTED PROBLEMS

CHAPTER 1

1.1 (a) 5 mA; (b) 5 kΩ; (c) 1 V; (d) 10 mA

1.3 (a) 5 V, 25 mW; (b) 5 kΩ, 5 mW; (c) 10 mA, 1 kΩ; (d) 10 V, 100 kΩ; (e) 31.6 mA, 31.6 V

1.5 990 kΩ, 190 kΩ, 90 kΩ, 10 kΩ; 9.9 kΩ (1% reduction), 9.09 kΩ (9.1% reduction), 5 kΩ (50% reduction)

1.7 2 V, 1.2 kΩ; 1.88 V to 2.12 V; 1.26 kΩ to 1.14 kΩ

1.9 4.80 V, Shunt the 10-kΩ resistor with 157 kΩ; Add a series resistance of 200 Ω.

1.11 10 kΩ, 5 kΩ

1.15 0.77 V, 12.31 kΩ, 0.05 mA

1.16 0.75 mA, 0.5 mA, 1.25 mA, 2.5 V

1.20 (a) 10^{-7} s, 10^7 Hz, 6.28×10^7 rad/s; (f) 10^3 rad/s, 1.59×10^2 Hz, 6.28×10^{-3} s

1.22 (a) $(1 - j\,1.59)$ kΩ; (b) $(247.3 - j\,1553)$ Ω; (c) $(71.72 - j\,45.04)$ kΩ; (d) $(100 + j\,628)$ Ω

1.24 60 mV, 1.2 μA, 50 kΩ

1.25 5 kΩ

1.29 (a) 165 V; (b) 24 V

1.32 14 kHz, 441 mV (peak); 312 mV; 693 mV, 71.4 μs

1.34 0, 110, 1011, 11100, 111011

1.36 (c) 12, 1.2 mV, 0.6 mV

1.38 7.056×10^5 bits/second.

1.40 11 V/V or 20.8 dB; 22 A/A or 26.8 dB; 242 W/W or 23.8 dB; 120 mW, 95.8 mW, 20.2%

1.43 (a) 82.6 V/V or 38.3 dB

1.46 0.69 V; 0.69 V/V or −3.2 dB; 8280 A/A or 78.4 dB; 5713 W/W or 37.6 dB.

1.48 S-A-B-L is preferred as it provides higher voltage gain.

1.51 (a) 400 V/V; (b) 40 kΩ, 2×10^4 A/A, 8×10^6 W/W; (c) 500 Ω; (d) 750 V/V; (e) 100 kΩ, 100 Ω, 484 V/V

1.56 38.1 V/V

1.59 Voltage amplifier, $R_i = 1 \times 10^5\ \Omega$, $R_o = 1 \times 10^2\ \Omega$, $A_{vo} = 121$ V/V

1.64 1025 V/V or 60.2 dB, 2500 A/A or 68 dB, 2.63×10^6 W/W or 64.2 dB

1.68 4 MHz

1.70 64 nF

1.73 $0.51/CR$

1.75 0.8 kΩ, 8.65 kΩ, connect 2 nF to node B.

1.78 159 kHz; 14.5 Hz; ≃ 159 kHz

1.81 10 Hz, 10 kHz, 0.04 dB, 0.04 dB, 10 Hz, 10 kHz

CHAPTER 2

2.2 4004 V/V

2.5 40,000 V/V

2.7 0.1%

2.8 In all cases, −5 V/V, 20 kΩ

2.11 (a) −1 V/V; (c) −0.1 V/V; (e) −10 V/V

2.13 10 kΩ, 100 kΩ

2.15 Average = +5 V, highest = +10 V, lowest = 0 V

2.17 $\pm 2x\%$; −98 to −102 V/V

2.19 1.8 kΩ; 18 kΩ

2.21 ±2 mV

2.24 $1000\left(1+\frac{R_2}{R_1}\right)$, $100\left(1+\frac{R_2}{R_1}\right)$, $10\left(1+\frac{R_2}{R_1}\right)$; $1000\,R_1$, $100\,R_1$, $10\,R_1$

2.26 (b) 1 kΩ, 100 kΩ, 909 V/V

2.28 (a) 10.2 kΩ

2.32 (a) 0.1 mA, 0.1 mA, 10 mA, 10.1 mA, −1 V; (b) 1.19 kΩ; (c) −11.1 V to −2.01 V

2.34 (a) 1 kΩ; (b) 0, ∞; (c) −0.57 mA to +0.57 mA (d) 2.2 mA

2.36 $v_O = -(10\,v_1 + 5\,v_2)$; −5 V

2.43 12.8 kΩ

2.44 (a) ∞; 0; (b) 10 kΩ, 10 kΩ; (d) 10 kΩ, 990 kΩ

2.46 100 kΩ; no

2.50 $2 \sin(2\pi \times 1000t)$

2.51 $1/x$; 1 to ∞; add 1-kΩ resistor between the left end of the pot and ground.

2.53 (a) 10 mV, 10 μA, 10 μA; (b) 10 V, 10 mA, 0; from the power supply of the op amp

2.58 (a) −0.83 V/V, 17%; (c) −0.98 V/V, 2%; (e) −9 V/V, 10%

2.60 20 V/V, 10 kΩ, 0.0095 V/V, 66.4 dB

2.64 $R_1 = R_3$, $R_2 = R_4$

2.65 0.02x V/V; 0.002 V/V, 54 dB; 0.02 V/V, 34 dB; 0.1 V/V, 20 dB

2.67 1 kΩ, 1 MΩ, 1 kΩ, 1 MΩ; 1% tolerances

2.69 $R = 1$ MΩ, $R_5 = 756$ Ω, $R_6 = 6.8$ kΩ

2.73 (a) −0.12 V to +0.12 V; (b) −12 V to +12 V

2.75 Ideal: 21 V/V, 0, ∞; ±1% resistors: $A_d = 21 \pm 4\%$, $|A_{cm}| = 0.02$, CMRR = 60.4 dB

2.77 (a) $v_B/v_A = 3$ V/V, $v_C/v_A = -3$ V/V; (b) 6 V/V; (c) 56 V pp, 19.8 V rms

2.79 (a) 1591 Hz; (b) leads by 90°; (c) increases by a factor of 10; (d) the same as in (b)

2.81 1 MHz; 0.159 μs

2.83 $R = 10$ kΩ, $C = 159$ pF; $R_F = 1$ MΩ, 1 kHz; (a) v_O decreases linearly to −6.3 V, (b) v_O decreases exponentially, $v_O(t) = -100(1 - e^{-t/159})$, reaching −6.1 V at the end of the pulse.

2.86 $R_1 = 10$ kΩ, $R_2 = 1$ MΩ, $C = 0.16$ nF; 100 kHz

2.88 15.9 kHz, $v_O = -5 \sin(10^6 t + 90°)$ V

2.90 Square wave of the same frequency, 8 V peak amplitude, average is 0 V; 30 kΩ

2.92 $R_1 = 1$ kΩ, $R_2 = 100$ kΩ, $C = 79$ nF; 20 Hz

2.94 4 mV

2.96 9 mV; 12 mV

2.98 (a) 0.53 μA, into the input terminals; (b) −3 mV; (c) −60 nA

2.100 $R_1 = 1.01$ kΩ, $R_2 = R_3 = 100$ kΩ, $C_1 = 1.58$ μF, $C_2 = 0.16$ μF

2.102 6 V; 3 V; 9 mV

2.104 (a) 0.2 V; (b) 0.4 V; (c) 10 kΩ, 20 mV; (d) 0.22 V

2.106 (a) 9.9 kΩ; (b) 0.222 V

2.108 80,000 V/V, 125 Hz, 10 MHz

2.111 19.61 kHz, 49.75 V/V, 4.975 V/V

2.113 (a) 5.1 MHz; (c) 10 MHz; (e) 10.1 MHz; (g) 2 MHz

2.116 36.6 MHz

2.118 500 MHz; 3; 7 MHz; 3.6 MHz

2.121 100 mV

2.125 1 MHz, 3.18 V

CHAPTER 3

3.1 -55°C: 2.68×10^6 cm^{-3}, one out of every 1.9×10^{16} silicon atoms; $+75^\circ$C: 3.70×10^{11} cm^{-3}, $N/n_i = 1.4 \times 10^{11}$

3.3 5×10^{18} cm^{-3}; 45 cm^{-3};

3.5 At 27°C: $n_n = 10^{17}$/cm^3, $p_n = 2.25 \times 10^3$ cm^3;
At 125°C: $n_n = 10^{17}$/cm^3, $p_n = 2.23 \times 10^8$ cm^3

3.7 $v_{p-\text{drift}} = 1.44 \times 10^6$ cm/s, $v_{n-\text{drift}} = 4.05 \times 10^6$ cm/s

3.9 9.26×10^{17}/cm^3

3.12 778 mV; 0.2 μm, 0.1 μm, 0.1 μm; 1.6×10^{-14}C

3.14 1.6 pC

3.16 59.6 mV

3.20 7.85×10^{-17} A; 0.3 mA

3.22 3.6×10^{-16} A; 0.742 V

3.24 31.6 fF; 14.16 fF

3.27 0.5 pF; 129.5 ps

CHAPTER 4

4.2 (a) -3 V, 0.6 mA; (b) $+3$ V, 0 mA

4.3 (a) $V = 2$ V, $I = 2.5$ mA; (b) $I = 1$ mA, $V = 1$ V

4.6 $X = AB$; $Y = A + B$; X and Y are the same for $A = B$; X and Y are opposite if A # B.

4.9 (a) $I = 0$, $V = 1$ V; (b) $I = 0.25$ mA, $V = 0$ V

4.11 $R \geq 4.2$ kΩ, 169.7 V

4.13 2.5 V; 1.25 V; 25 mA; 12.5 mA; 2.5 V

4.15 34 V; 8.3 Ω; 0.6 A; 29 V; 34 V, 8 Ω; 25%; 103 mA; 0.625 A; 29 V

4.17 At -55°C, $V_T = 18.8$ mV; At $+55^\circ$C, $V_T = 28.3$ mV; $V_T = 25$ mV at 17° C.

4.19 0.335 μA

4.21 (a) 6.91×10^{-13} A, 0.64 V; (c) 5.11×10^{-17}A, 0.59 V

4.23 3.9 mA; -22 mV

4.26 $A_4 = 2A_3 = 4A_2 = 8A_1$; 1.5 mA

4.28 42 Ω

4.31 50°C; 6 W; 8.33°C/W

4.33 230 mV; independent of temperature

4.35 0.6635 V, 0.3365 mA

4.37 $R = 582\ \Omega$

4.41 (a) −2.3 V, 0.53 mA; (b) +3 V, 0 mA

4.43 (a) $I = 0$, $V = -1.23$ V; (b) $I = 0.133$ mA, $V = 0$ V

4.45 $R \geq 4.23$ kΩ, 169.7 V; essentially the same.

4.48 0.24 mV, 2.0 mV, 9.6 mV; 25 μA

4.53 $V_o/V_i = 1/(1 + j\omega C r_d)$; $-\tan^{-1}(\omega C\, V_T/I)$; 157 μA −84.3° to −5.71°

4.56 $R = 417\ \Omega$; 7.39 mA; 6.8 mV; −3.4 mV; −6.8 mV; −13.6 mV

4.59 (a) $r_z = 8\ \Omega$, 1.04 W; (b) $V_{Z0} = 8.8$ V, 188 mW

4.61 88.9 mV

4.63 167 Ω; 7.65 V; 7.35 V; 7.78 V; 707 Ω; 7.2 V

4.66 (a) 9.825 V; (b) 207 Ω; (c) 33 mV/V, ±1.65%; (d) −6.77 V/A, −1.35%; (e) 70.9 mA, 732 mW

4.69 0.324 V

4.71 13.44 V; 48.4%; 8.3 V; 8.3 mA

4.73 (a) 10.1:1; (b) 1.072:1

4.75 45 V

4.77 (a) 12.77 V, 13.37 V; (b) 7.1%, 2.24%; (c) 192 mA, 607 mA; (d) 371 mA; 1.2 A

4.80 (a) 9.7 V; (b) 542 μF; (c) 25.7 V, 38.5 V; (d) 739 mA; (e) 1.42 A

4.83 10.74 V; 23.5 μs; 4.913 A; 4.913 A

4.85 (a) +1 V, +2 V, +2.7 V; (b) +3 V, +6 V, +6.7 V; (c) 0 V, 0 V, 0 V, −13 V; (d) 0 V, 0 V, −13 V.

4.96 −7.07 V

CHAPTER 5

5.1 580 to 2900 μm^2; 24 to 54 μm

5.4 (a) 0.5; (b) 0.5; (c) 1.0; (d) 0.5

5.7 0.35 μm

5.9 0.5 V; 0.5 mA

5.11 (a) −1.1 V; (b) −0.4 V; (c) 0.05 mA; 0.5 mA

5.13 116.3 Ω, 116.3 mV; 50

5.17 2.8 V; 500 Ω, 100 Ω

5.19 5 mA/V^2; 0.6 V

5.21 0.5 V; 20; 145 μA; 1.5 V, 1.125 mA

5.23 2.5 kΩ to 125 Ω; (a) 5 kΩ to 250 Ω; (b) 1.25 kΩ to 62.5 Ω; (c) 2.5 kΩ to 125 Ω.

5.29 (a) 3%; (b) 5%

5.31 200 kΩ, 20 kΩ; 5%, 5%

5.33 104 μA; 4%; double L to 3 μm

5.35 Increases by a factor of 16.

5.38 350 μA; 750 μA; 864 μA; 880 μA; 960 μA

5.41 At 3.0 V, transistor is cut off; at 2.5 V, transistor enters saturation region; at 0.5 V, transistor enters triode region.

5.43 1 V, 0 V, 1 V, 0.25 V; 5 kΩ, 5 kΩ, 5 kΩ, 5 kΩ; 10 kΩ, 2 V; 10 kΩ, −1 V; 10 kΩ, 2 V; 10 kΩ, −0.75 V

5.45 0.08 mA; 10 kΩ, 5 kΩ; 17.5 kΩ

5.48 4 kΩ

5.50 4 μm, 11.1 μm; 1.4 kΩ

5.52 0.45 mA, +7.3 V; quite tolerant.

5.54 44.4; 1.25 kΩ

5.56 −1 V, −1.43 V, −2.8 V, 1 V, 2.8 V, +1 V, 2.8 V, −1 V

5.59 $I_1 = 405$ μA, $V_2 = 1.5$ V; $I_3 = 217$ μA, $V_4 = 1.232$ V; $V_5 = 1.5$ V, $I_6 = 405$ μA

5.61 (a) 0.5 V, 0.5 V, −0.983 V; (b) 0.1 V, 0.9 V, −1.01 V

5.63 −1.24 V

5.65 triode, 0.59 mA; triode, 5 mA; saturation, 9 mA; saturation, 9 mA

CHAPTER 6

6.2 4.7×10^{-17} A, 1.87×10^{-16}A; $A_2/A_1 = 4$

6.4 0.31 V

6.6 Old: 0.673 V; New: 0.846 V

6.8 80; 0.988

6.10 0.5; 0.67; 0.91; 0.95; 0.98; 0.99; 0.995; 0.998; 0.999

6.12 $I_C = 0.5$ mA → 3 mA; $I_E = 0.51$ mA → 3.01 mA; 30 mW

6.14 990 μA, 99, 0.99; 980 μA, 49, 0.98; 950 μA, 19, 0.95

6.17 −0.668 V; 1.04 V; 0.02 mA

6.19 EBJ: 0.691 V; CBJ: 0.576 V; EBJ: 0.49 μA; CBJ: 48.5 μA

6.23 0.758 V; 0.815 V

6.25 238 mA; 6×10^{-14} A; 87

6.28 (a) 2 mA, −0.7 V; (b) −2 V; (c) 2 V, 0.5 mA; (d) 1.6 mA, −4.5 V

6.30 8.3 kΩ; 20; 100; 200

6.32 $R_C = 4$ kΩ; $R_E = 3.64$ kΩ; $R_{C\max} = 5.86$ kΩ

6.34 $R_E = 3.66$ kΩ; $R_C = 5$ kΩ

6.36 10.24 μA

6.38 0.75 V; 0.55 V

6.40 3.35 μA; 3000

6.43 125 kΩ; 125 V; 12.5 kΩ

6.45 1 mA; 10 V; 50 V; 50 kΩ

6.47 $\beta = 100$; $\beta_{ac} = 80$; $\Delta i_C = 0.18$ mA, $i_C = 1.18$ mA

6.50 $\beta_{forced} = 11.2$; $V_C = 4.8$ V; $R_B = 45.7$ kΩ

6.52 2.05 V, 2.38 V

6.55 $R_1 = 18$ kΩ, $R_2 = 12$ kΩ; 0.46 mA, 2.54 V

6.58 +0.41 V, +1.11 V, −1.15 V; +1.2 V, +1.9 V, −1.9 V; 204

6.61 (a) −0.7 V, +1.2 V; (b) +1.2 V, 0.5 mA; (c) −0.7 V, 0 V, +1.2 V; (d) +1.45 V, −0.5 V; (e) +0.75 V, +1.45 V, −0.5 V

6.63 $R_E = 4$ kΩ, $R_B = 50$ kΩ, $R_C = 4$ kΩ, $I_C = 0.85$ mA to 0.98 mA, $V_C = -1.6$ V to −1.1 V

6.66 (a) 0 V, +0.7 V, −0.725 V, −1.425 V, +1.1 V; (b) +0.23 V +0.93 V, −1 V, −1.7 V, +1.47 V

6.68 0 V, 0 V; +1.8 V, +1.1 V; −2.2 V, −1.5 V; −3 V, −2.3 V

CHAPTER 7

7.2 A: (0.5 V, 5 V); B: (0.72 V, 0.22 V)

7.3 20 kΩ; (0.72 V, 0.22 V); −40 V/V; 0.78 V; 19.5 mV

7.6 0.4 V; 8.33

7.8 (a) 0.712 V; (b) −42.7 V/V, 11.7 mV; (c) 42.88 kΩ; 24.9

7.10 −160 V/V; 0.7 V; 4.4 mV

7.12 1.08 V; 0.78 V; −156.7 V/V

7.15 −60 V/V

7.18 3 mA; −120 V/V; +5 mV: exp. → −660 mV, linear → −600 mV;
−5 mV: exp. → +540 mV, linear → +600 mV.

7.25 (a) 0.1 mA, 0.8 V; (b) 1 mA/V; (c) −10 V/V; (d) 100 kΩ, −9.1 V/V

7.26 0.5 mA/V; 0.067 mA, 0.27 V; 9.14; 0.67 V

7.29 16 μm; 0.75 V

7.31 −18.2 V/V; 1.207 V, −23.6 V/V

7.33 (b) 2 mA/V, 200 kΩ; (d) 3.33 MΩ, 0.94 V/V, −15.38 V/V, −14.5 V/V

7.35 2.5 V; 0.611 mA, 1.95 V; 5 mV; −0.55 V; −110 V/V; −100 V/V

7.37 40 mA/V; 25 Ω; 2.5 kΩ; 1 V

7.39 1.04 kΩ to 4.7 kΩ

7.42 (a) 1.000, ∞, 1.00 mA, 1.00 mA, 0 mA, 40 mA/V, 25 Ω, ∞ Ω; (c) 0.980, 50, 1.00 mA, 1.02 mA, 0.02 mA, 40 mA/V, 24.5 Ω, 1.25 kΩ; (e) 0.990, 100, 0.248 mA, 0.25 mA, 0.002 mA, 9.92 mA/V, 100 Ω, 10.1 kΩ

7.48 1 V; 125 Ω; 80 V/V

7.53 $R_{in} = 75\ \Omega$; $v_o/v_{sig} = 39.6$ V/V

7.55 −1000 V/V; −5000 V/V

7.57 8.6 kΩ, 7.7 kΩ; 77 V/V

7.59 79.4 V/V; 4762 A/A

7.64 −10 V/V

7.66 1 mA/V; 125 μA; −7.5 V/V

7.68 5 kΩ, 10 kΩ, −200 V/V; −100 V/V, −33.3 V/V; 15 mV, 0.5 V

7.70 (b) 1250 V/V

7.72 0.5 kΩ

7.74 30.3 kΩ, −40 V/V, 12 kΩ; −20 V/V, −15 V/V; 6.65 mV, 100 mV

7.76 80 V/V, 44.4 V/V to 109.1 V/V; $R_e = 275\ \Omega$, 25 V/V, 20 V/V to 27.3 V/V

7.78 2.5 mA/V; 0.2 V

7.80 $i_{sig}R_C$

7.82 0.357 kΩ; 1.6 mA; 1.13 V

7.84 1.25 mA; 1.5 mA, 1.0 mA; 0.5 V/V; 1 V

7.86 149 Ω, 0.87 V/V; 116 Ω to 246 Ω; 0.80 V/V to 0.90 V/V

7.89 −91 V/V

7.91 27.5 V/V, 41.2 V/V, 55.6 V/V, 57.1 V/V, 55.6 V/V; 0.325 mA

7.92 18 MΩ, 22 MΩ, 3 kΩ, 3 kΩ; 2 V

7.94 5.07 V, 1.27 mA to 2.48 mA; 620 Ω; 0.91 mA to 1.5 mA

7.96 2 V; 2.4 V; 1.2 mA

7.101 (a) 2.7 V, 2.2 V; (b) 3.05 V, 3.05 V

7.103 2.5 kΩ, 22 MΩ, 20 MΩ

7.105 (a) 230 kΩ; 0.5 mA to 1.5 mA; 1 V to 0 V (saturated transistor), design very intolerant of β variation.

7.108 (a) 5.73; (b) $V_{BB} = V_{BE} + 0.352\ V_{CC}$; (c) 38.8 kΩ, 37.5 kΩ, 3.33 kΩ; (d) 8.1 kΩ; 0.475 mA to 0.509 mA with a nominal value of 0.5 mA

7.110 5.75 kΩ, 6.2 kΩ; 10.8%

7.112 (a) $R_C = 1.5$ kΩ, $R_B = 80$ kΩ; (b) $R_C = 1.5$ kΩ, $R_B = 82$ kΩ; 1.52 V, 0.98 mA; (c) 0.7 V, 1.53 mA; (d) $R_{B1} = 40$ kΩ, $R_{B2} = 70$ kΩ, $R_C = 1.47$ kΩ, 1.1 V, 1.28 mA

7.116 8.6 kΩ, +0.4 V

7.118 (a) $V_D = 2.5$ V, $k_n = 11.1$ mA/V^2; (b) 120 kΩ, −4.1 V/V; (c) 0.264 V, 1.08 V; (d) 300 Ω, 1.08 V

7.120 20 mA/V; 0.1 mA; 5 mV; 10 kΩ

7.122 (a) 0.99 V/V, 99 Ω; (b) 99 Ω, 14.3 V/V; (c) 7.15 V/V

7.124 (a) 1.6 V, 0.1 mA, 82.4 kΩ; (b) 1 mA/V; (d) 1.95 V/V, 39.1 kΩ

7.126 $R_1 = 47$ kΩ, $R_2 = 24$ kΩ, $R_E = 2.2$ kΩ, R_C either 4.7 kΩ or 5.1 kΩ

7.128 $R_B = 91$ kΩ, $R_C = 22$ kΩ, $I = 0.2$ mA; −176 V/V; −29.7 V/V

7.130 (a) 1 mA, 8.2 V; (c) 2.32 kΩ, 0.32 V/V; (d) 2.32 kΩ, −69.2 V/V; (e) −61.8 V/V; (f) 1368.5 V/V

7.132 (a) 0.495 mA, 1.18 V; −71.9 V/V

7.134 $\beta = 50$: (a) 0.78 mA, 0.78 V, 1.48 V; (b) 21.3 kΩ; (c) 0.64 V/V; $\beta = 200$: (a) 1.54 mA, 1.54 V, 2.24 V; (b) 50.9 kΩ; (c) 0.81 V/V

7.136 (a) 1.73 mA, 68.4 mA/V, 14.5 Ω, 1.4645 kΩ; (b) 148.3 kΩ, 0.93 V/V; (c) 18.21 kΩ, 0.64 V/V

7.138 75 Ω; 25 Hz; 25 V/V

CHAPTER 8

8.1 12 kΩ; 0.2 V; 25 kΩ; 20 μA

8.3 50; 8.75 kΩ

8.6 5 μm, 25 μm, 10 μm, 2.5 μm, 5 μm; 15 kΩ; 25 kΩ, 31.25 kΩ

8.8 (a) 0.691 V to 0.863 V, 10 μA to 10 mA; (b) 9.62 μA, 0.098 mA, 0.98 mA, 9.62 mA

8.11 0.1 mA, 10%

8.14 Both cases: −0.7 V, +2 V, +0.7 V, −0.7 V, −1.7 V; (a) $I = 0.4$ mA; (b) $I = 0.04$ mA

8.17 700 Ω, 5 A/A, 10 kΩ.

8.19 $v_o = g_{m1} v_i \, (W_3/W_2) R_L$; $g_{m1} R_L \, (W_3/W_2)$; $1/g_{m2}$; $-g_{m1}/g_{m2}$

8.21 (a) 1.6 kΩ; (b) 250 Ω

8.25 $I = 10$ μA: 0.4 mA/V, 250 kΩ, 1 MΩ, 400 V/V; $I = 100$ μA: 4 mA/V, 25 kΩ, 100 kΩ, 400 V/V; $I = 1$ mA: 40 mA/V, 2.5 kΩ, 10 kΩ, 400 V/V

8.27 50 V/V; 0.2 mA; 12.5 μm

8.29 0.4 μm; 25; 0.2 mA

8.31 0.5 mA; 4 mA/V

8.33 1 mA/V; 15 kΩ; 15 V/V; 3.9 μm

8.35 0.144 mA

8.37 (a) 80 μA/V, 0.18 MΩ, 14.4 V/V; (b) 0.79 V, 0.253 mA/V, 18 kΩ, 4.55 V/V; (c) 0.8 mA/V, 18 kΩ, 14.4 V; (d) 0.08 V, 0.253 mA/V, 180 kΩ, 45.5 V/V; (e) lowest A_0: first design when operated at $I_D = 100$ μA, $A_0 = 4.55$ V/V, highest A_0: second design when operated at $I_D = 10$ μA, $A_0 = 45.5$ V/V; gain increases by a factor 10.

8.39 0.5 μm; 12.5

8.41 1.05 V; 2 μm; 8; 32

8.43 (a) 0.95 V, 0.475 μA, 2.375 V; (b) −86.5 V/V, 1.9 V, 22 mV; (c) 33.7 kΩ

8.45 0.913 V; 1.07 V

8.47 (a) 25 μA; (b) 0.33 V and 2.98 V; (c) −189.3 V/V; (d)−195.8 V/V; (e) −210.6 V/V

8.49 (a) 0.25 mA; (b) 120 kΩ, 120 kΩ, 60 kΩ; (c) 5 kΩ, 10 mA/V; (d) 5 kΩ, −600 V/V, 60 kΩ

8.51 980 Ω; 61 kΩ; 10.1 V/V

8.53 2 kΩ; 1.1 V

8.55 (a) 100 μA, 1.03 V; (b) 0.9 mA/V, 200 kΩ; (c) 2.2 kΩ; (d) 209 kΩ; (e) 90.9 V/V, 89 V/V; (f) 32 mV

8.57 r_o

8.59 0.99 (or more exactly, 0.975); 14.8 MΩ

8.61 (a) 208 Ω; (b) 500 Ω; (c) 4.8 kΩ; 101 with $R_e = \infty$

8.63 (a) 50, 1.6 MΩ; (b) 250, 320 kΩ

8.65 0.5 μm; 20; 1 V; 0.25 mA; 0.5 V

8.67 0.6 μm; 0.125 mA; $(W/L)_{1,2} = 10$; $(W/L)_{3,4} = 40$

8.69 $g_{m2} r_{o2}$

8.71 0.2 V; 0.5 V to 0.8 V

8.74 1.2 V; 1.0 V; 0.8 V; 100; 6.91 MΩ

8.76 1 MΩ

8.79 -10^5 V/V

8.81 (a) 1.41 mA/V, 822.3 kΩ, −1159 V/V; (b) 1.41 mA/V, 457 kΩ, −644 V/V

8.83 $(g_{m3}r_{o3})\ (g_{m2}r_{o2})r_{o1}$

8.85 (a) $I_{O1} = I_{O2} = \frac{1}{2}I_{REF}/(1+2/\beta^2)$; (b) Use $I_{REF} = 0.7$ mA and 3 transistors Q_3, Q_4 and Q_5 whose EBJ areas are in the ratio 1:2:4; currents realized are 0.0999 mA, 0.1999 mA and 0.3997 mA.

8.88 (a) 0.3 V, 0.8 V; (b) 8 μA, 172 μA; (c) 180 μA; (d) 1.1 V; (e) 12 MΩ; (f) 0.08 μA, 0.04%

8.90 (a) $R_E = 2.88$ kΩ; (b) 8.2 MΩ, 0.7 μA

8.92 (a) 58.5 kΩ; (b) 79.9 MΩ,

8.95 360 μA; 2.4 mA/V; 0.48 mA/V; 27.8 kΩ; 0.81 V/V; 339 Ω; 0.7 V/V

8.97 (b) $g_{m1} = 0.632$ mA/V, $g_{m2} = 40$ mA/V, $r_{\pi 2} = 5$ kΩ; (c) −19.5 V/V; (d) 487 kΩ, −9.6 V/V; (e) 10 MΩ, −18.6 V/V

8.99 50.2 V/V

CHAPTER 9

9.1 (a) 0.2 V, 0.6 V; (b) −0.6 V, 0.08 mA, 0.08 mA, +0.6 V, +0.6 V; (c) −0.2 V, 0.08 mA, 0.08 mA, +0.6 V, +0.6 V; (d) −0.7 V, 0.08 mA, 0.08 mA, +0.6 V, +0.6 V; (e) 1.0 V; (f) −0.8 V, −0.2 V

9.3 (a) 0 V, −0.6 V, 0.6 V, 0.6 V, 0 V; (b) 0.104 V, −0.541 V, 0.4 V, 0.8 V, 0.4 V; (c) 0.283 V, −0.4 V, +0.2 V, 1 V, 0.8 V; (d) −0.104 V, −0.645 V, +0.8 V, +0.4 V, −0.4 V; (e) −0.283 V, −0.683 V, +1 V, +0.2 V, −0.8 V

9.5 0.587 V; −0.587 V; 0.612 V; 0.025 V; 0.10 V, 4 V/V; −0.025 V

9.7 0.35 V; 16.3; 1.14 mA/V

9.9 0.212 V; 554.5 μA

9.11 (a) 0.1 V_{OV}; (b) 0, 0.338 V_{OV}, 0.05 V_{OV}, 0.005 V_{OV}, 1.072 V_{OV}

9.13 0.25 V; 0.5 mA; 5 kΩ; 40

9.15 0.5 mA; 3.6 kΩ; 38.6

9.17 $I = 2I_D$; $P_{diff} = 2P_{CS}$

9.19 (a) $g_{m1,2}\left[\frac{1}{g_{m3,4}} \| r_{o3,4} \| r_{o1,2}\right]$; (b) $\sqrt{[\mu_n(W/L)_{1,2}]/[\mu_p(W/L)_{3,4}]}$; (c) 25

9.23 8 kΩ; W/L, I_D (mA) and $|V_{GS}|$ (V) are: Q_1(50, 0.1, 0.7), Q_2(50, 0.1, 0.7), Q_3(100, 0.2, 0.7), Q_4(20, 0.1, 0.7), Q_5(20, 0.1, 0.7), Q_6(100, 0.2, 0.7), Q_7(40, 0.2, 0.7)

9.25 0.632 μm; 0.28 mA

9.27 $v_{B1} = +0.5$ V: −0.177 V, +0.52 V, 2.5 V; $v_{B1} = -0.5$ V: −0.677 V, +2.5 V, +0.52 V

9.30 (a) −0.574 V, 0.4 V, 0.4 V; (b) −0.326 V to 0.674 V; (c) 5 mV

9.32 (a) $V_{CC} - (I/2)R_C$; (b) 2 V; (c) 0.4 mA, 5 kΩ

9.34 $R_C = 5.05$ kΩ, +1.6 V

9.36 0.5 mA, 1.0 mA; 17.3 mV

9.38 8 mA/V; 40 kΩ

9.40 5 mV; 250 Ω; −40 V/V; 200 mV; 400 mV

9.42 Each emitter has a resistance $R_e = 450$ Ω, $R_C = 10$ kΩ; $I = 1$ mA; Possible value of $V_{CC} = 10$ V

9.49 12 V/V

9.51 16 V/V

9.53 25 V/V; 101 kΩ

9.55 7.7 V/V; 5×10^{-4} V/V; 1.54×10^{4} or 83.8 dB

9.57 (a) 2.332 V; (b) 5.06 kΩ; (c) 2.47 V; (d) −1.92 V/V; (e) 0.287 V

9.59 4 μm

9.61 (a) 20 V/V; (b) 0.23 V/V; (c) 86.5 or 38.7 dB; (d) $-0.023 \sin 2\pi \times 60t + 0.2 \sin 2\pi \times 1000t$, volts

9.63 (a) 100 V/V; (b) 50 kΩ; (c) 2.5×10^{-4} V/V; (d) 4×10^{5} or 112 dB; (e) 25 MΩ

9.65 (a) 50 V/V; (b) 2.5×10^{-3} V/V, 2×10^{4} or 86 dB; (c) 5×10^{-5} V/V, 10^{6} or 120 dB

9.67 (a) Two emitter resistances and a single bias-current source I; $R_e = 25$ Ω; $R_C = 10$ kΩ; $V_{CC} = +15$ V; $R_{EE} = 50$ kΩ; $V_A = 100$ V; 2.4 MΩ

9.69 2/3 in one transistor and 1/3 in the other; 0.008 V/V

9.72 11 mV; variability of V_t; 7.33%

9.74 2.5 mV

9.77 −0.25 mV

9.79 1.25 mV

9.81 (a) $x = 0.3$ kΩ; (b) $x = 0.225$ kΩ

9.83 $2\alpha I/3$ and $\alpha I/3$; $\alpha IR_C/3$; 18.75 mV; 17.3 mV

9.85 20 kΩ; 40 V/V

9.87 1.4 mA/V; 25 kΩ; 25 kΩ; 17.5 V/V

9.89 3 V

9.92 1 mA/V; 75 kΩ; 75 V/V; 75 kΩ

9.94 20 kΩ; 20 kΩ; 10 mA/V; 200 V/V; 100 V/V

9.96 $-2V_T/\beta_P^2$; $-20\ \mu V$

9.98 2.67×10^4 V/V

9.100 $\frac{I/2}{\beta+1}\bigg/\left(\frac{\beta}{2}\right)$, a reduction by a factor of $(\beta/2)$; R_{id} increases by a factor $(\beta/3)$

9.102 1.13 mA/V; 75 kΩ; 85 V/V

9.105 1 mA/V; 25 kΩ; 25 V/V; 25 kΩ, 0.02 mA/V; 0.98 kΩ; 0.98 A/A; 50 kΩ; 2600 kΩ; −0.0196 V/V; 1274 or 62.1 dB

9.107 0.1

9.110 8 mA/V; 100 kΩ; 800 V/V; 37.5 kΩ; 100 kΩ; −0.013 V/V; 60,000 or 96 dB; 444.4 V/V

9.112 (a) 83.3 kΩ; (b) 1200 V/V; (c) 21×10^6 or 146 dB

9.114 (a) W/L: 12.5, 12.5, 50, 50, 25, 100, 25, 25, 0 V; (b) −0.1 V to +0.7 V; (c) −0.7 V to +0.7 V; (d) 900 V/V

9.116 108 μA; 909 mV; 0.86 mV

9.118 (a) W/L: 32.9, 32.9, 178, 178, 65.8, 356, 65.8, 32.9; (b) 0.65 V to 1.05 V; (c) 0.15 V to 1.05 V; 144 V/V

9.120 25 V/V; 20 kΩ; 5000 A/A

9.122 R_5; 7.37 kΩ; reduced to about half its original value; change R_4 to 1.085 kΩ, this will slightly reduce A_2.

9.124 (a) 0.52 mA, 1.04 mA, 2.1 mA, 0 V; (b) 4 kΩ, 65.5 Ω; (e) 8770 V/V

CHAPTER 10

10.1 20 nf

10.3 10 μF; 88.4 Hz; 8.84 Hz

10.5 (a) 10 kΩ; (b) 3.53 μF; (c) 10 Hz; (d) 100 Hz; (e) dc gain = 2, makes perfect sense since C_S behaves as an open-circuit at dc.

10.7 5 μF; 0.5 μF; 0.5 μF; 92.2 Hz; 6 μF

10.10 141.4

10.13 $g_m = 1.3$ mA/V; $g_{mb} = 0.25$ mA/V; $r_o = 100$ kΩ; $C_{gs} = 61.6$ fF; $C_{gd} = 4.3$ fF; $C_{sb} = 12.8$ fF; $C_{db} = 9.4$ fF; $f_T = 3.1$ GHz

10.17 $L = L_{min}$: 6.5 V/V, 113 GHz; $2L_{min}$: 13 V/V, 28.3 GHz; $3L_{min}$: 19.5 V/V, 12.6 GHz; $4L_{min}$: 26 V/V, 7.1 GHz; $5L_{min}$: 32.5 V/V, 4.5 GHz

10.19 265.3 MHz

10.21 500 MHz; 600 MHz; 252 ps; 0.43 pF

10.23 50 MHz; 10 MHz

10.25 5 pF; < 31.8 kΩ

10.28 200.2 pF; $-1000/[1 + sC_{in}R_{sig}]$; 795 kHz; 795 MHz

10.31 870 kHz; −6.1 V/V; $R_{in} = 33.3$ kΩ → 3.1 V/V; $R_L = 1.24$ kΩ → 1.6 V/V

10.33 −9.2 V/V; 525 kHz

10.35 61 pF; 522 kHz

10.37 −33 V/V; 873 kHz; 28.8 MHz; f_H increases by a factor of 1.16 and voltage gain decreases by the same factor while GB remains nearly constant. Power dissipation increases by a factor of 2.

10.39 −32.8 V/V; 572 kHz

10.41 (a) 1001 pF, 1.001 pF; (c) 20 pF, 20 pF; (e) −90 pF, 9 pF; +90 pF

10.44 (a) 0.54 mA; (b) 21.6, A/V, 4.63 kΩ; (c) −10.8 V/V; (d) 4 kΩ, 2.14 kΩ; (e) −7.4 V/V; (f) 14.37 pF; (g) 16.3 MHz

10.46 −80 V/V; 3.8 MHz; 6.4 GHz; 304 MHz

10.48 −81.4 V/V; 21.4 MHz; 11.2 GHz

10.50 (a) 99.2 MHz; (b) 227.6 MHz

10.53 (a) 4.26; (b) 49.3

10.55 5.67×10^7 rad/s

10.57 −40.6 V/V; $\tau_{gs} = 243.8$ ns; $\tau_{gd} = 3112.8$ ns; $\tau_{CL} = 300$ ns; 43.5 MHz

10.59 −80 V/V; 10.1 pF; 788 kHz; 652 kHz; the latter as it takes into account C_L.

10.61 41.6 fF

10.63 −138.9 V/V; 2.98 MHz; 2.28 MHz, the latter as it takes into account C_L.

10.66 8.3 V/V; 239 MHz; 7.23 MHz; 7.23 MHz

10.69 11.1 fF

10.71 −913 V/V; 6.28 MHz

10.73 0.2 V; 0.2 mA; 289.4 MHz; 57.9 kHz, −100 V/V (40 dB)

10.76 −26.5 V/V; 5.7 MHz

10.78 −100,000 V/V; 31.8 kHz, 31.8 kHz; 3.18 GHz

10.79 0.91 V/V; 200 Ω; 398 MHz; 33.4 MHz, 90.7 MHz; 31.6 MHz

10.82 $0.8/[s^2 + 8.886 \times 10^6\ s + 39.48 \times 10^{12}]$

10.84 0.96 V/V; 2 GHz; 676 MHz, 4.6 GHz; 676 MHz

10.86 1.59 MHz

10.88 4 MHz; decreases by a factor of 4 to 1 MHz

10.90 (b) −49.8 V/V; (c) 53.2 pF, 598 kHz, 29.8 MHz

10.92 50 V/V; 15.9 MHz; 1.59 GHz; 3.18 GHz

10.96 (a) −100 V/V, 603 kHz, 60.3 MHz; (b) −50 V/V, 1.02 MHz, 51.2 MHz

10.101 (a) 2.5 MΩ, −4000 V/V; (b) 107.6 kHz; two dominant capacitances: C_L (most significant) and $C_{\mu 2}$

10.103 66.7 V/V; 2 MHz

10.106 (a) 10,000 V/V; (b) 11.1 MHz

CHAPTER 11

11.1 4.9×10^{-3}; 169.5; −15.3%

11.3 1; 0.999 V/V; 60 dB; 0.999 V, 0.001 V; −0.011%

11.5 (b) (i) 1000; (ii) 100; (iii) 20

11.7 2500 V/V; 0.0196 V/V; 49; 50 V/V; 34 dB

11.10 99; 4

11.12 1000 V/V; 0.099 V/V

11.14 416.6 V/V; 9.33×10^{-3} V/V; 5016.8 V/V, 9.95×10^{-3} V/V; 41.66 V/V, 9.33×10^{-2} V/V; 501.68 V/V, 9.95×10^{-2} V/V

11.16 500 V/V; 0.098 V/V; 653.4 V/V

11.19 1 MHz, 1 Hz

11.21 Three stages; each with a closed-loop gain of 10 V/V, an amount of feedback of 100, and $\beta = 0.099$ V/V.

11.23 50 V/V; 0.008 V/V; 16 Hz

11.25 0.089; for $|v_S| \leq 0.9$ V, $v_O/v_S = 11.1$ V/V, for 0.9 V $\leq |v_S| \leq 1.4$ V, $v_O/v_S =$ 10.1 V/V, and for $|v_S| \geq 1.4$ V, $v_O = \pm 15$ V

11.27 (a) 90 kΩ; (b) 43.11, 9.77 V/V; (c) 2.343

11.29 (a) $1 + \dfrac{R_2}{R_1} = 11$ V/V; (b) 0.1 mA, 0.3 mA, +7.7 V; (c) 23.2; (d) 10.5 V/V

11.31 (a) 0.9 kΩ; (b) 31.33, 9.7 V/V, −3%, change R_F to 933 Ω

11.33 (a) 47.62 β, 47.62 V/V; (b) 821 kΩ, 179 kΩ

11.35 Lower; 199; 20 kΩ

11.37 100 V/V; 1.001 MΩ

11.39 (a) $1 + (R_2/R_1) = 11$ V/V; (b) 0.1 mA, 0.3 mA, +7.7 V; (c) 255.3 V/V, 0.359 kΩ, 0.917 kΩ; (d) 1/11; (e) 10.5 V/V, 8.59 kΩ, 39.4 Ω, 4.5%

11.41 (b) 10 V/V; (c) 0.2 V, 1.1 V, 0.2 V, 0.9 V; (d) −35.3 V/V, −50 V/V, 0.935 V/V, 1650 V/V; (e) 0.1 V/V; (f) 9.94 V/V, −0.6%; (g) 5.6 Ω

11.44 (c) 1.2 kΩ; (d) 1.42 kΩ, 628 Ω; (e) 23.8 V/V; (f) 145 kΩ, 0.53 Ω

11.46 100 Ω; 9.94 mA/V

11.48 (c) −0.999 kΩ

11.50 (a) 0.135 V/V; (b) 7.4 V/V; (c) 0.14 Ω

11.53 (a) 200 Ω; (b) 1418.4 mA/V; (c) 283.7, 284.7; (d) 4.982 mA/V, very close; (e) 28.2 kΩ, 8 MΩ

11.56 9.56 mA/V; 503.4 kΩ

11.58 (a) 0 V, +0.6 V, +0.6 V; (b) 0.1 mA/V; (c) 0.099 mA/V; (d) 203 MΩ; (e) 0.99 V/V; 1.25 Ω

11.60 −9.88 kΩ, 11.1 Ω, 1.1 Ω compared to −9.99 kΩ, 1.11 Ω, 0.11 Ω.

11.62 3.23; −0.1 mA/V; −32.3 kΩ; −7.63 kΩ; due to the approximation used in the systematic analysis method.

11.64 (a) $-R_F/R_s$, 20 kΩ; (b) −9.88 V/V, 21.7 Ω, 22.1 Ω; (c) 82.18 kHz

11.66 159, larger by about 2.5%, a result of the approximations involved in the general method. The more accurate value is the one obtained here.

11.68 10 kΩ; −9.52 kΩ; 11.9 Ω; 244 Ω

11.70 (b) −98.8 V/V; 7.2 Ω; 10.3 Ω

11.72 0.53 kΩ; 10.5 Ω; 526 Ω

11.74 (d) −99.8 A/A, −0.1 A/A, 9.98, −9.1 A/A, 0.2 kΩ, 18.2 Ω; (e) 328.4 kΩ

11.76 970.9, −9709 A/A, −9.99 A/A; $A\beta$ and A differ slightly from the results in Example 11.10; however, A_f is identical.

11.81 $I_{C1} = 0.1$ mA, $I_{C2} = 10$ mA; $V_o/V_s = 3.62$ V/V; $R_{in} = 176.7$ Ω

11.83 20 krad/s; 4×10^{-3} V/V; 250 V/V

11.85 8×10^{-4}

11.87 10 V/V; 10^5 Hz; 1 MHz; by the amount-of-feedback $\simeq 10^4$.

11.89 (a) 2.025×10^{-4}, 5.5×10^4 Hz; (b) 3306 V/V, 1653 V/V; (c) 0.5; (d) $(-5.5 \pm j\,13.25) \times 10^4$ Hz, 1.325

11.91 0.1; 0.686; 2.1

11.93 2; 173.2 kHz

11.95 3.085×10^3 Hz; 18.15°; 10^{-3}

11.97 3.16×10^{-4}; 2.4×10^3 V/V or 67.6 dB.

11.99 2.4×10^4 V/V or 87.6 dB; 9.09×10^3 V/V or 79.2 dB.

11.101 2 kHz; 500

11.104 10 Hz; 15.9 nF

11.106 (b) 3.16×10^4 Hz, 1.8°; (c) zero: -10^3 rad/s, poles: $(-0.505 \pm j\,31.62) \times 10^3$ rad/s, the response is very peaky with a peak of 1000 at 31.62 krad/s.

CHAPTER 12

12.1 −9.3 V to +9.7 V; −8.6 V to +10.4 V; −4.65 V to +9.7 V; −3.95 V to +10.4 V; −9.7 V to +9.7 V; −9 V to +10.4 V

12.3 2.7 kΩ; 24 mW

12.6 $V_{CC}I$ (in all cases)

12.8 $\hat{V}$; $\hat{V}/R_L$; 25%

12.11 4.5 V; 6.4%; 625 Ω

12.13 10 V; 6.37 V; 6.85 Ω, 7.3 W; 9.62 Ω, 1.3 W

12.17 1.266 V; 12.5 Ω; 0.889 V/V; 0.998 V/V

12.19 2.15 mA

12.22 1 mA; −1.06 V; +4 V; −6 V

12.24 0.98 mA; +5.1 V; −10 V; 99; 1.96 mA; 1.92 mA

12.28 20.7 mA; 788 mW; 7.9°C; I_Q becomes 37.6 mA, etc., etc.

12.30 (a) 1.365 kΩ, 1.365 kΩ, 1.365 V; (b) 1.420; (c) 1.512 V; (d) 1.641 V

12.32 (a) For $R_L = \infty$: at $v_I = 0$, $I_1 = 0$; at $v_I = +10$ V, $I_1 = 20$ μA, at $v_I = -10$ V, $I_1 = -20$ μA; (b) $R_L = 100$ Ω; at $v_I = 0$, $I_1 = 0$, at $v_I = +10$ V, $I_1 = 22.5$ μA, at $v_I = -10$ V, $I_1 = -22.5$ μA.

12.34 215 Ω, 215 Ω, 0.75 Ω, 0.75 Ω; 0.7 Ω; 0.704 V

12.37 (a) 0.0164 mA, 1.64 mA; (b) 32.8 v_i, −66.2 V/V; (c) 27.2 kΩ

12.39 $R_1 = 300$ kΩ, $R_2 = 632$ kΩ; 9.484 V and −10.644 V

12.41 3.84 Ω; 384 mV; 0.94 μA

12.43 6.5 Ω; 487.5 mV; 2.9 μA

12.45 (b) 1.25 V, 1.56 mA

12.47 (a) Q_1 : 35.6, Q_2 : 88.9, Q_N : 356, Q_P : 889; (b) −0.6 V; (c) 1.38 V

12.49 ±2.05 V

12.51 (b) 0.15 V

12.53 (a) 533.3, 1333.3; (b) 10 V/V; (c) −5%; (d) 1.85 V and −1.85 V; (e) 0.3 V and −0.3 V; (f) −1.77 V to +1.77 V

12.55 R_2 and R_3; R_3; R_2; $R_2 = 33.3$ kΩ and $R_3 = 1.33$ kΩ

12.57 16 V; 2.7 W; 13 V $p-p$

12.59 30 kΩ, 40 kΩ

12.62 +3 V; −3 V

12.64 (c) 8 Ω, 5 A, 50 W; (d) 6 Ω, 5 A, 37.5 W; (e) 3 Ω, 5A, 18.75 W

12.66 12.5°C/W; 8 W; 112.5°C

12.68 (a) 37.5°C/W; (b) 1.33 W; (c) 62.5°C

12.70 72°C; 1.5°C/W; 4 cm

CHAPTER 13

13.1 −0.8 V to +1.2 V; −0.8 V to +0.8 V

13.3 0.15 V

13.5 0.45 μm; 2000 V/V

13.7 (a) 10,000 V/V; (b) 10^8 rad/s and 10^7 rad/s; (c) 10^9 rad/s, 4 pF, 25×10^3 rad/s, 5×10^8 rad/s

13.9 (a) 1.59 pF; (b) $f_{P1} = f_t/A_0$, $f_{P2} = 318$ MHz, $f_Z = 200$ MHz; (c) 46°; (d) 500 Ω, 72.5°; (e) 722 Ω

13.11 125.6 V/μs; 0.8 pF

13.13 (a) 2 pF; (b) 1.51 pF

13.15 (a) 0.16 V; (b) 2 pF; (c) 78.1

13.17 (b) 0.45 μm

13.19 250 μA; 400 μA; 200 μA; 50 μA

13.21 25, 25, 25, 25, 6.25, 6.25, 6.25, 6.25, 125, 125, 50

13.23 100 μA; 150 μA; 15.92 MHz; 54.7°; 6.58 MHz; $C_L = 24.2$ pF; 4.13 V/μs

13.25 0.12 V; $I_B = I = 150$ μA; 15 V/μs; W/L : 26, 26, 65, 65, 26, 26, 26, 26, 130, 130, 52

13.28 (a) −0.25 V to +1.3 V; (b) −1.3 V to +0.25 V; (c) −0.25 V to +0.25 V; (d) −1.3 V to +1.3 V

13.30 $C_p = 0.176\, C_L$

13.33 $V_{EB} = 625$ mV; A device: 7.3 mA/V, 137 Ω, 6.85 kΩ, 278 kΩ; B device: 21.9 mA/V, 46 Ω, 2.28 kΩ, 90.9 kΩ

13.35 $I_3 = I_1 \left\{ \left[\frac{1}{\sqrt{k_1}} + \frac{1}{\sqrt{k_2}} \right] \Big/ \left[\frac{1}{\sqrt{k_3}} + \frac{1}{\sqrt{k_4}} \right] \right\}^2$; 0.1 mA

13.37 603 mV; 518 mV; 8.5 kΩ

13.39 4.75 μA; $R_4 = 1.94$ kΩ

13.41 14 μA

13.43 53.3 nA; 20.1 nA

13.45 −3 V to +4.8 V

13.47 6.4 kΩ; 270 μA

13.49 1.68 mA; 50.4 mW

13.51 4.63 kΩ

13.53 10 mV

13.55 0.691 μA; 3.6 mV

13.57 $R = 18.2$ kΩ; 15.55 MΩ

13.60 3.1 MΩ, 9.38 mA/V

13.62 −3.6 V to +4.2 V

13.64 14.4 Ω

13.66 20.2 mA; double the value of R_7

13.68 5.67 MHz

13.70 180 Hz; 0.7 pF

13.73 159.2 kHz; 10^8 rad/s or 15.9 MHz

13.75 (a) 0.05 mA, 0.05 mA, 0.05 mA, 0.05 mA, 1 mA, 1 mA, 1 mA; (b) 100 kΩ; (c) 5×10^4 V/V or 94 dB; (d) 63.7 pF

13.77 $Q_5 : Q_1 = 1$; $Q_6 : Q_1 = 4$; 3.47 kΩ; 3 MΩ and 7 MΩ

13.79 (a) 0.1 V to 2.2 V; (b) 0.8 V to 2.9 V

13.81 12.5 kΩ; 0.8 V to 3.35 V; 100 kΩ; 10 μA, 50 kΩ

13.83 $36.9/I$; 1240 V/V; $1240(IR_L)/(IR_L + 36.9)$; 5.1 μA, 11.8 μA

13.85 (a) 0.1 V to 2.9 V; (b) 20 kΩ; (c) 0.2 Ω; (d) 12.3 mA, 0.3 mA, 1.6 kΩ; (e) 0.3 mA, 12.3 mA, 2.4 kΩ

13.88 10.6 μA; 0.3 mA

CHAPTER 14

14.1 (a) 2.18 kΩ; (b) 5.40 kΩ; (c) 3.71

14.2 (a) 6; (b) 1.67 kΩ

14.16 0.6 V; 0.7 V

14.18 $NM_H = 0.2\ V_{DD}$; $NM_L = 0.3\ V_{DD}$; $0.2\ V_{DD}$; 2 V

14.20 (a) 0.12 V, 2.5 V, 1.5 V, 0.68 V; (b) $V_{OH} = 2.5 - 0.4N$, $NM_H = 1.5 - 0.4N$, $N = 2$; (c) (i) 3 mW, (ii) 1 mW

14.22 $V_{IL} = 0.776$ V, $V_{IH} = 0.816$ V; $NM_H = 1.184$ V; $NM_L = 0.776$ V; -50 V/V

14.24 $V_{DD} = 1.2$ V, $R_D = 38.3$ kΩ, $W/L = 1.5$; 0 W, 36 μW

14.26 $V_{DD} = 1.2$ V, $R_D = 23$ kΩ, $W/L = 2.5$; 0.435 V, 0.6 V, 0.7 V, 0.385 V, 0.5 V

14.29 6.84

14.31 (a) 244 nm, 22,181 nm^2; (b) 1 V, 0 V, 0.5375 V, 0.4625 V, 0.4625 V, 0.4625 V (c) both equal; 2.18 kΩ

14.33 1.82

14.35 40.1

14.37 (a) 0.78 μm, 0.127 μm^2; (b) 1.3 V, 0 V, 0.7125 V, 0.5875 V, 0.59 V, 0.59 V, 0.0625 V, 1.24 V, 0.53 V, 0.53 V; (c) 1.48 kΩ, 1.48 kΩ; (d) -5.8 V/V, 0.762 V, 0.538 V, 0.224 V; (e) 0.57 V, -0.08 V, 60%; (f) 0.61 V, -0.04 V, 40%

14.39 (a) $v_O(t) = 10\, e^{-t/\tau}$; (b) 69 ns, 220 ns

14.41 69 ps, 35 ps, 52 ps

14.43 (a) 1.2 ns, 0.6 ns; (b) 1 pF; (c) $C_{\text{out}} = 0.6$ pF, $C_{\text{load}} = 0.4$ pF

14.45 30 ps, 60 ps, 45 ps

14.47 57.5 ps, 69 ps, 63.3 ps

14.49 $(W/L)_n \geq 1.725$, $(W/L)_p \geq 4.14$

14.51 34.4 ps, 42.6 ps, 38.5 ps; 13 GHz

14.53 36.3 ps, 36.3 ps, 36.3 ps; 9.35 fF

14.55 (c) $14.66 \times 10^3\ (2C_n + C_w)$; (d) $8.625 \times 10^3\ (3.4C_n + C_w)$ (e) (i) In both cases, $t_P = 29.32 \times 10^3 C_n$, thus when C is entirely intrinsic, scaling does not affect t_P; (ii) For $W_p = W_n$, $t_P = 14.66 \times 10^3 C_w$, and for $W_p = 2.4W_n$, $t_P = 8.625 \times 10^3 C_w$, thus using a matched design reduces t_P only when C is dominated by external capacitance.

14.60 (a) 2.65 V; (b) 2.24 V

14.63 32.4 fJ; 64.8 W; 36 A

14.65 0.36 pF

14.67 32 pJ

14.69 (a) t_P and the maximum operating frequency remain unchanged, *PDP* is reduced by a factor of 0.52; (b) t_P increases by a factor (1/0.72) and the maximum operating frequency is reduced by the factor of 0.72. The *PDP* decreases by a factor of 0.72.

CHAPTER 15

15.1 4.88×10^8 or 488 million transistors

15.3 260 cm^2/Vs, 144.4 cm^2/Vs; E_{cr}(NMOS) $= 3.85 \times 10^4$ V/cm; E_{cr}(PMOS) $= 6.92 \times 10^4$ V/cm

15.5 (b) 0.62

15.7 (b) 2.75

15.9 (a) 207 pA; (b) 207 mA, 207 mW

15.11 (a) 270 Ω; (b) 0.1 pF; (c) 93.2 ps

15.13 1.3 V; 0.095 V; 40.5 μA; 52.7 μW

15.15 167 ps; 36.9 ps; 102 ps

15.17 2.1; 0.5 V; 0.5 V, 0.47 V, 0.44 V

15.19 1.69; 0.58 V; 152 μW

15.23 1.26

15.26 0.834 V

15.28 25.8 ps

15.30 2.07 V, 0 V; 10.4 μA; 0.9 ns; 0.5 ns

15.34 13.5 μA; 351.6 μA; 182.6 μA; 0.18 ns

15.36 (a) 1.2 V, 0 V; (b) 240 μA, 60 μA, 7.8 μA, 56.25 μA, 49.4 ps; (c) 240 μA, 60 μA, 225 μA, 1.9 μA, 34.2 ps, 0.466 V; 41.8 ps

15.39 8.3 kΩ; 83 ps

15.45 0.188 ns

15.47 0.188 ns; 0.077 ns

15.49 (d) 0.35 V, 0.6 V

15.51 2 GHz

15.53 −1.453 V, −1.205 V, −1.73 V, −0.88 V; 0.230 V, 0.325 V, 0.345 V

15.55 22.45 mW

15.57 1 V; +5 V; $(A+B).(C+D)$

15.59 2.6 V; 8.18 mA

CHAPTER 16

16.1 A(0 V, 0 V), B(2.5 V, 2.5 V), C(5 V, 5 V); 25 V/V; 0.2 V

16.4 $(W/L)_{1,3} = 0.13$ μm/0.13 μm, $(W/L)_{2,4} = 0.52$ μm/0.13 μm, $(W/L)_{5,8} =$ 0.26 μm/0.13 μm

16.6 $(W/L)_{5,6} = 3.83$, higher than the values without velocity saturation to compensate for the current reduction resulting from velocity saturation.

16.7 0.4 μm/0.13 μm; 65 ps

16.11 4,294,967,296

16.13 16

16.15 57%

16.17 $(W/L)_a \leq 4.5$

16.19 4.5; (i) 0.23 V, 121.8 μA; (ii) 0.34 V, 158.7 μA; (iii) 0.4 V, 180 μA

16.22 1.75, greater than the value without velocity saturation because of the current reduction due to velocity saturation.

16.24 (a) 3; (b) 4.93 ns; (c) 3.33 ns

16.26 3

16.29 $L = 0.13\ \mu\text{m}$, $(W/L)_n = (W/L)_p = (W/L)_a = 1$

16.31 128 Mbits

16.33 0.5 pA

16.35 0.4 mA/V; 353 mV; 130 mV; 100% (doubling); 4 ns

16.37 $(W/L)_n = 3.33$, $(W/L)_p = 13.32$; 1.44 ns; 2 ns

16.39 (a) 0.4 V; (b) 0.1 V, 0.3 V; (c) 132 μA; (d) $(W/L)_{1,2} = 26.4$, $(W/L)_{3,4} = 6.6$, $(W/L)_5 = 52.8$

16.41 10; 1024; 10,240; 1024; 12,288

16.43 40 MHz, 48%

16.45 4

16.48 (a) 2.4 ns; (b) 22 ns, 3.16 V; (c) 1.9 ns

CHAPTER 17

17.2 (a) 0.995 V, −5.7°; (b) 0.707 V, −45°; (c) 0.1 V, −84.3°; (d) 0.01 V, −89.4°

17.4 1 V/V; 0.977 V/V; 0.001 V/V

17.6 0.97 dB; 14.15 dB

17.10 (a) LP: $T(s) = 10^{20}/(s+10^4)(s^2+0.618\times10^4 s+10^8)(s^2+1.618\times10^4 s+10^8)$ (b) HP: $T(s) = s^5/(s+10^4)(s^2+0.618\times10^4 s+10^8)(s^2+1.618\times10^4 s+10^8)$;

17.12 $T(s) = 0.2656\,(s^2+4)/(s^2+0.5s+1.0625)$; 0.2656

17.14 $1/(s^3+2s^2+3s+2)$; -1, $-0.5 \pm j1.323$

17.17 35.7 dB

17.19 $N = 4$; $2\pi\times10^4(-0.383 \pm j0.924)$, $2\pi\times10^4\,(-0.924 \pm j0.383)$; $\omega_0^4/(s^2+0.765\,\omega_0 s+\omega_0^2)\times(s^2+1.848\,\omega_0 s+\omega_0^2)$ where $\omega_0 = 2\pi\times10^4$ rad/s; 38.2 dB

17.22 0.975 rad/s, 0.782 rad/s, 0.434 rad/s, 0 rad/s; 1 rad/s, 0.901 rad/s, 0.623 rad/s, 0.223 rad/s; −64.9 dB; 42 dB/octave

17.24 (a) $N = 10$, 4 dB; (b) Normalized to $\omega_p = 2\pi \times 3.4 \times 10^4$ rad/s, the poles are: $-0.0224 \pm j0.9978$; $-0.0651 \pm j0.9001$; $-0.1013 \pm j0.7143$; $-0.1277 \pm j0.4586$; $-0.1415 \pm j0.1580$, $T(s) = 7.60 \times 10^4/(s^2 + s\ 0.0448\ \omega_p + 0.9961\ \omega_p^2)(s^2 + 0.1302\ \omega_p + 0.8144\ \omega_p^2)\ (s^2 + s\ 0.2026\ \omega_p + 0.5205\ \omega_p^2)(s^2 + 0.2554\ \omega_p + 0.2266\ \omega_p^2)(s^2 + s\ 0.2830\ \omega_p + 0.0450\ \omega_p^2)$

17.26 $R_1 = 120$ kΩ; $C = 6.63$ nF; $R_2 = 120$ kΩ

17.28 $R_1 = 10$ kΩ, $R_2 = 10$ kΩ, $C_1 = 0.16$ μF, $C_2 = 1.6$ nF; High-frequency gain= 40 dB

17.30 $T(s) = -(s - \omega_0)/(s + \omega_0)$ where $\omega_0 = 1/CR$; $T(j\omega) = \left(1 - j\dfrac{\omega}{\omega_0}\right) \Big/ \left(1 + j\dfrac{\omega}{\omega_0}\right)$; $-2\tan^{-1}(\omega/\omega_0)$; 5.36 kΩ, 11.55 kΩ, 20 kΩ, 34.60 kΩ, 74.63 kΩ.

17.33 $T(s) = 10^8/(s^2 + 5000\,s + 10^8)$; 9.354 krad/s, 2.066

17.35 $T(s) = s^2/(s^2 + \sqrt{2}\,s + 1)$; Zeros: two at $s = 0$; Poles: $-0.707 \pm j0.707$

17.37 $T(s) = \pi \times 10^4 s/[s^2 + \pi \times 10^3 s + (2\pi \times 10^4)^2]$; Zeros: $s = 0$ and $s = \infty$; Poles: $1.57 \times 10^3 \times (-1 \pm j39.988)$

17.39 $[s^2 + (2\pi \times 60)^2]/[s^2 + s(2\pi \times 60) + (2\pi \times 60)^2]$

17.42 $T(s) = (1/LC)/[s^2 + s/CR + (1/LC)]$

17.44 (a) -0.5%; (b) -0.5%; (c) no change

17.46 $s^2 \Big/ \left(s^2 + \dfrac{1}{CR} + \dfrac{1}{LC}\right)$

17.49 $V_o = \left[s^2 V_y + s\left(\dfrac{\omega_0}{Q}\right)V_z + \omega_0^2 V_x)\right] \Big/ \left[s^2 + s\left(\dfrac{\omega_0}{Q}\right) + \omega_0^2\right]$

17.51 $R_1 = R_2 = R_3 = 10$ kΩ; (a) $C_4 = 0.15$ μF; (b) $C_4 = 15$ nF; (c) $C_4 = 1.5$ nF

17.55 First-order section (Fig. 17.13a): $R_1 = R_2 = 100$ kΩ, $C = 10$ nF; Second-order section (Fig. 17.22a): $C_4 = C_6 = 10$ nF, $R_1 = R_2 = R_3 = R_5 = 100$ kΩ, $R_6 = 161.8$ kΩ, $K = 1$; Second-order section (Fig. 17.22a): $C_4 = C_6 = 10$ nF, $R_1 = R_2 = R_3 = R_5 = 100$ kΩ, $R_6 = 61.8$ kΩ, $K = 1$

17.57 $C_4 = C_6 = 1$ nF, $R_1 = R_2 = R_3 = R_5 = 79.6$ kΩ, $R_6 = 159.2$ kΩ, $r_1 = r_2 = 10$ kΩ

17.60 (b) First-order section: $C = 1$ nF, $R_1 = R_2 = 13.71$ kΩ, Second-order LPN section: $R_1 = R_2 = R_3 = R_5 = 9.76$ kΩ, $C_{61} = 618$ pF, $C_{62} = 382$ pF, $R_6 = 35.9$ kΩ, $K = 1$

17.62 (b) $C = 1$ nF, $R = 10$ kΩ, $R_1 = 10$ kΩ, $R_f = 10$ kΩ, $R_2 = 10$ kΩ, $R_3 = 70$ kΩ, $R_L = R_H = 10$ kΩ, $R_B = 40$ kΩ, $R_F = 57.1$ kΩ

17.64 1%

17.67 (b) First-order section: $C = 1$ nF, $R_1 = R_2 = 13.71$ kΩ, Second-order LPN section: $C = 1$ nF, $R = 9.76$ kΩ, $R_d = 35.9$ kΩ, $r = 10$ kΩ, $C_1 = 618$ pF, $R_1 = R_3 = \infty$, $R_2 = 9.76$ kΩ

17.71 $\omega_0 = 6/CR$, $Q = 3$, Center-frequency gain= -18 V/V.

17.73 (a) Q^2; (b) $2Q^2$

17.75 (b) Second-order section [Fig. 17.34(c)]: $R_1 = R_2 = 10\text{ k}\Omega$, $C_3 = 492\text{ pF}$, $C_4 = 5.15\text{ nF}$;
Second-order section [Fig. 17.34(c)]: $R_1 = R_2 = 10\text{ k}\Omega$, $C_3 = 1.29\text{ nF}$, $C_4 = 1.97\text{ nF}$;
First-order section (Fig. 17.13a): $R_1 = R_2 = 10\text{ k}\Omega$, $C = 1.59\text{ nF}$

17.77 $S_L^{\omega_0} = -\frac{1}{2}, S_C^{\omega_0} = -\frac{1}{2}, S_R^{\omega_0} = 0; S_L^Q = -\frac{1}{2}, S_C^Q = \frac{1}{2}, S_R^Q = 1$

17.79 $S_A^{\omega_0} = 0, S_A^Q = 2Q^2/A$

17.81 $S_{C_4}^{\omega_0} = S_{C_6}^{\omega_0} = S_{R_1}^{\omega_0} = S_{R_3}^{\omega_0} = S_{R_5}^{\omega_0} = -\frac{1}{2},\ S_{R_2}^{\omega_0} = +\frac{1}{2},\ S_{R_6}^Q = +1,\ S_{C_6}^Q = S_{R_2}^Q = +\frac{1}{2},$
$S_{C_4}^Q = S_{R_1,\,R_2,\,R_3}^Q = -\frac{1}{2},$

17.83 1 mA/V; 0.99 kΩ

17.85 0.314 mA/V

17.87 $G_{m1} = 2.51$ mA/V; $G_{m2} = 0.251$ mA/V

17.90 $C_1 = Q^2 C$; $G_m = \omega_0 QC$

17.92 $G_m = 0.785$ mA/V; $G_{m2} = 0.785$ mA/V; $G_{m3} = 0.157$ mA/V; $G_{m4} = 0.785$ mA/V

17.94 1 pC; 0.1 μA; 0.1 V; 100 cycles; 10^4 V/s

17.96 $C_3 = C_4 = 6.283$ pF; $C_5 = 0.126$ pF; $C_6 = 0.126$ pF

17.98 80.3 rad/s; 83; 967 kHz; 66.7 V/V

17.100 838.8 kHz; 47.4

17.103 A (dB): 7, 8.5, 9.3, 9.8, 10.1; W/B: 31.6, 8.6, 5.9, 4.9, 4.5

CHAPTER 18

18.1 ω_0; $AK = 1$

18.3 (a) 1; (b) 2

18.5 0.6 mA/V; 15.92 MHz

18.7 120°; $\omega_0 = \sqrt{3}/CR$; $2/R$

18.11 $\omega_0 = 1/CR$; $Q = 1/3$; Gain$= 1/3$

18.13 $\omega_0 = 1/CR$; $Q = 1\Big/\left(2 - \frac{R_2}{R_1}\right)$

18.15 $\omega_0 = 1/CR$; $R_2/R_1 \geq 2$

18.17 7.88 V

18.19 $f_0 = 406$ Hz; $R_f = 290\text{ k}\Omega$

18.22 9.95 kΩ; 3.6 V; add a diode in series with each of the limiter diodes.

18.24 $\omega_0 = 1\Big/\sqrt{L\left(\frac{C_1 C_2}{C_1 + C_2}\right)}$; simplified condition: $g_m R_L > \frac{C_2}{C_1}$

18.26 $\omega_0 = 1\Big/\sqrt{L\left(\frac{C_1C_2}{C_1+C_2}\right)}$; $g_m R_L' > \frac{C_1}{C_2}$

18.28 (b) $\omega_0 = 1/\sqrt{LC}$; $IR_C > 0.1$ V, (c) $(4/\pi)$ V

18.30 2.0165 MHz to 2.0173 MHz, a range of 800 Hz.

18.32 (a) $V_{TH} = \left(\frac{L_+}{R_2} + \frac{V}{R_3}\right)(R_1 \| R_2 \| R_3)$; $V_{TL} = \left(\frac{L_-}{R_2} + \frac{V}{R_3}\right)(R_1 \| R_2 \| R_3)$;
(b) $R_2 = 656.7$ kΩ, $R_3 = 19.7$ kΩ

18.36 (a) Output will be either +12 V or −12 V; (b) The output is a symmetric square wave (±12 V) of frequency f and it lags the sine wave by an angle of 65.4°; 0.1 V.

18.38 1989 Hz

18.40 $V_Z = 3.6$ V; $R_1 = R = 25$ kΩ; $R_3 = 5.83$ kΩ; $C = 0.01$ μF; $R = 25$ kΩ

18.42 96 μs

18.44 $C_1 = 1$ nF, $C_2 = 0.1$ nF, $R_1 = R_2 = 100$ kΩ, $R_3 = 134.1$ kΩ, $R_4 = 470$ kΩ; 5.8 V; 61.8 μs

18.46 (a) 18.2 kΩ; (b) 10.67 V

18.48 (b) 100.6 kHz, 75%; (c) 15.6 μs, 55.2 kHz, 86.2%; 3.90 μs, 156 kHz, 61%

18.50 1.85 V

INDEX

References marked **P** are study Problems; **S** are points in the chapter Summaries; Appendix pages found in the book or on the website are shown as **B17**; and Additional Material found on the website is **AM1**.

C

E

F

G

H

I

J

K

L

N

O

P

Q

R

S

U

V

W

X

Y

Z

9.7 DATA CONVERTERS—AN INTRODUCTION

In this section we begin the study of another group of analog IC circuits of great importance; namely, data converters.

9.7.1 Digital Processing of Signals

Most physical signals, such as those obtained at transducer outputs, exist in analog form. Some of the processing required on these signals is most conveniently performed in an analog fashion. For instance, in instrumentation systems it is quite common to use a high-input-impedance, high-gain, high-CMRR differential amplifier right at the output of the transducer. This is usually followed by a filter whose purpose is to eliminate interference. However, further signal processing is usually required, which can range from simply obtaining a measurement of signal strength to performing some algebraic manipulations on this and related signals to obtain the value of a particular system parameter of interest, as is usually the case in systems intended to provide a complex control function. Another example of signal processing can be found in the common need for transmission of signals to a remote receiver.

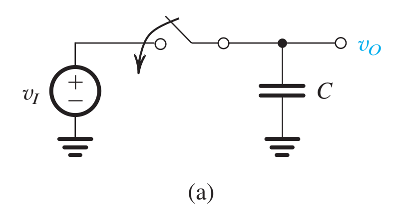
Many such forms of signal processing can be performed by analog means. In earlier chapters we encountered circuits for implementing a number of such tasks. However, an attractive alternative exists: It is to convert, following some initial analog processing, the signal from analog to digital form and then use economical, accurate, and convenient digital ICs to perform **digital signal processing.** Such processing can in its simplest form provide us with a measure of the signal strength as an easy-to-read number (consider, e.g., the digital voltmeter). In more involved cases the digital signal processor can perform a variety of arithmetic and logic operations that implement a **filtering algorithm.** The resulting **digital filter** does many of the same tasks that an analog filter performs—namely, eliminate interference and noise. Yet another example of digital signal processing is found in digital communications systems, where signals are transmitted as a sequence of binary pulses, with the obvious advantage that corruption of the amplitudes of these pulses by noise is, to a large extent, of no consequence.

Once digital signal processing has been performed, we might be content to display the result in digital form, such as a printed list of numbers. Alternatively, we might require an analog output. Such is the case in a telecommunications system, where the usual output may be audible speech. If such an analog output is desired, then obviously we need to convert the digital signal back to an analog form.

It is not our purpose here to study the techniques of digital signal processing. Rather, we shall examine the interface circuits between the analog and digital domains. Specifically, we shall study the basic techniques and circuits employed to convert an analog signal to digital form **(analog-to-digital** or simply **A/D conversion)** and those used to convert a digital signal to analog form **(digital-to-analog** or simply **D/A conversion).** Digital circuits are studied in Chapters 10 and 11.

9.7.2 Sampling of Analog Signals

The principle underlying digital signal processing is that of **sampling** the analog signal. Figure 9.36 illustrates in a conceptual form the process of obtaining samples of an analog signal. The switch shown closes periodically under the control of a periodic pulse signal (clock). The closure time of the switch, τ, is relatively short, and the samples obtained are



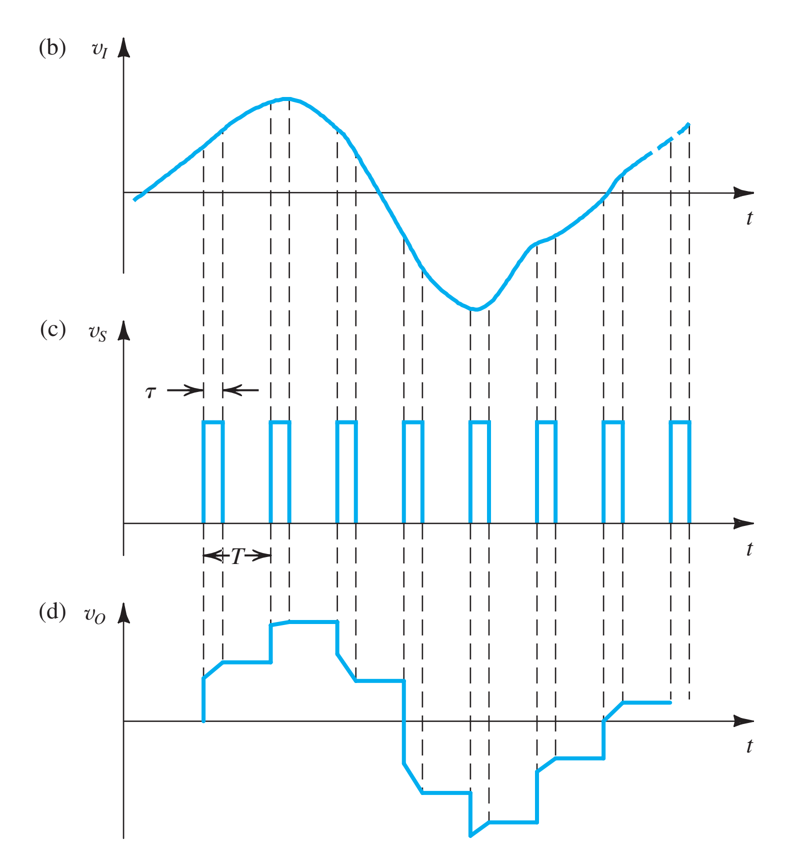


FIGURE 9.36 The process of periodically sampling an analog signal. **(a)** Sample-and-hold (S/H) circuit. The switch closes for a small part (τ seconds) of every clock period (T). **(b)** Input signal waveform. **(c)** Sampling signal (control signal for the switch). **(d)** Output signal (to be fed to A/D converter).

stored (held) on the capacitor. The circuit of Fig. 9.36 is known as a **sample-and-hold (S/H) circuit.** As indicated, the S/H circuit consists of an analog switch that can be implemented by a MOSFET transmission gate (Section 10.5), a storage capacitor, and (not shown) a buffer amplifier.

Between the sampling intervals—that is, during the *hold* intervals—the voltage level on the capacitor represents the signal samples we are after. Each of these voltage levels is then fed to the input of an A/D converter, which provides an N-bit binary number proportional to the value of signal sample.

The fact that we can do our processing on a limited number of samples of an analog signal while ignoring the analog-signal details between samples is based on the Shannon's sampling theorem [see Lathi (1965)].

9.7.3 Signal Quantization

Consider an analog signal whose values range from 0 to +10 V. Let us assume that we wish to convert this signal to digital form and that the required output is a 4-bit digital signal.[4] We know that a 4-bit binary number can represent 16 different values, 0 to 15; it follows that the *resolution* of our conversion will be 10 V/15 = $\frac{2}{3}$ V. Thus an analog signal of 0 V will be represented by 0000, $\frac{2}{3}$ V will be represented by 0001, 6 V will be represented by 1001, and 10 V will be represented by 1111.

All these sample numbers are multiples of the basic increment ($\frac{2}{3}$ V). A question now arises regarding the conversion of numbers that fall between these successive incremental levels. For instance, consider the case of a 6.2-V analog level. This falls between 18/3 and 20/3. However, since it is closer to 18/3 we treat it as if it were 6 V and *code* it as 1001. This process is called **quantization.** Obviously errors are inherent in this process; such errors are called quantization errors. Using more bits to represent (encode or, simply, code) an analog signal reduces quantization errors but requires more complex circuitry.

9.7.4 The A/D and D/A Converters as Functional Blocks

Figure 9.37 depicts the functional block representations of A/D and D/A converters. As indicated, the **A/D converter** (also called an **ADC**) accepts an analog sample v_A and produces an N-bit **digital word.** Conversely, the **D/A converter** (also called a **DAC**) accepts an n-bit digital word and produces an analog sample. The output samples of the D/A converter are often fed to a sample-and-hold circuit. At the output of the S/H circuit a staircase waveform, such as that in Fig. 9.38, is obtained. The staircase waveform can then be smoothed by a

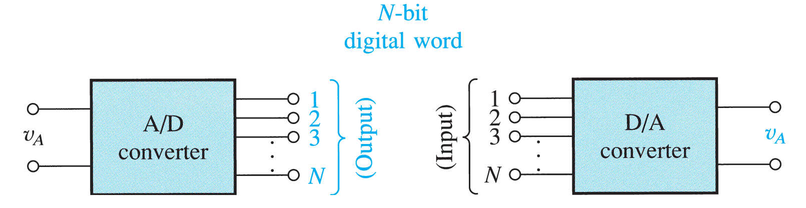


FIGURE 9.37 The A/D and D/A converters as circuit blocks.

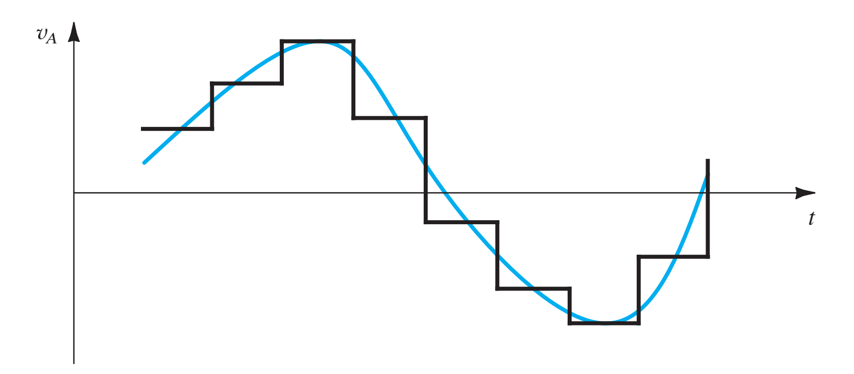


FIGURE 9.38 The analog samples at the output of a D/A converter are usually fed to a sample-and-hold circuit to obtain the staircase waveform shown. This waveform can then be filtered to obtain the smooth waveform, shown in color. The time delay usually introduced by the filter is not shown.

---

[4] *Bit* stands for *b*inary dig*it*.

low-pass filter, giving rise to the smooth curve shown in color in Fig. 9.38. In this way an analog output signal is reconstructed. Finally, note that the quantization error of an A/D converter is equivalent to $\pm\frac{1}{2}$ least significant bit (b_N).

EXERCISE

9.31 An analog signal in the range 0 to +10 V is to be converted to an 8-bit digital signal. What is the resolution of the conversion in volts? What is the digital representation of an input of 6 V? What is the representation of an input of 6.2 V? What is the error made in the quantization of 6.2 V in absolute terms and as a percentage of the input? As a percentage of full scale? What is the largest possible quantization error as a percentage of full scale?

Ans. 0.0392 V; 10011001; 10011110; –0.0064 V; –0.1%; –0.064%; 0.196%



9.8 D/A CONVERTER CIRCUITS

9.8.1 Basic Circuit Using Binary-Weighted Resistors

Figure 9.39 shows a simple circuit for an N-bit D/A converter. The circuit consists of a reference voltage V_{REF}, N binary-weighted resistors R, $2R$, $4R$, $8R$, . . . , $2^{N-1}R$, N single-pole double-throw switches S_1, S_2, . . . , S_N, and an op amp together with its feedback resistance $R_f = R/2$.

The switches are controlled by an N-bit digital input word D,

$$D = \frac{b_1}{2^1} + \frac{b_2}{2^2} + \cdots + \frac{b_N}{2^N} \tag{9.109}$$

where b_1, b_2, and so on are bit coefficients that are either 1 or 0. Note that the bit b_N is the **least significant bit (LSB)** and b_1 is the **most significant bit (MSB).** In the circuit in Fig. 9.39, b_1 controls switch S_1, b_2 controls S_2, and so on. When b_i is 0, switch S_i is in position 1, and when b_i is 1 switch S_i is in position 2.

Since position 1 of all switches is ground and position 2 is virtual ground, the current through each resistor remains constant. Each switch simply controls where its corresponding current goes: to ground (when the corresponding bit is 0) or to virtual ground (when the corresponding bit is 1). The currents flowing into the virtual ground add up, and the sum flows

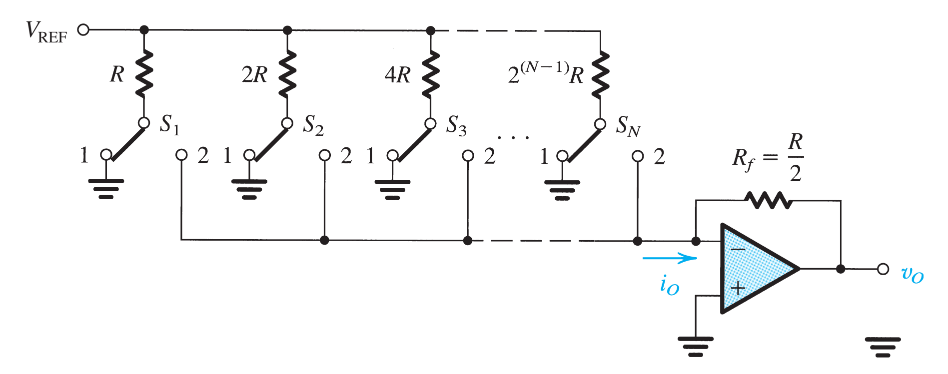


FIGURE 9.39 An N-bit D/A converter using a binary-weighted resistive ladder network.

through the feedback resistance R_f. The total current i_O is therefore given by

$$i_O = \frac{V_{\text{REF}}}{R}b_1 + \frac{V_{\text{REF}}}{2R}b_2 + \cdots + \frac{V_{\text{REF}}}{2^{N-1}R}b_N$$
$$= \frac{2V_{\text{REF}}}{R}\left(\frac{b_1}{2^1} + \frac{b_2}{2^2} + \cdots + \frac{b_N}{2^N}\right)$$

Thus,

$$i_O = \frac{2V_{\text{REF}}}{R}D \tag{9.110}$$

and the output voltage v_O is given by

$$v_O = -i_O R_f = -V_{\text{REF}}D \tag{9.111}$$

which is directly proportional to the digital word D, as desired.

It should be noted that the accuracy of the DAC depends critically on (1) the accuracy of V_{REF}, (2) the precision of the binary-weighted resistors, and (3) the perfection of the switches. Regarding the third point, we should emphasize that these switches handle analog signals; thus their perfection is of considerable interest. While the offset voltage and the finite on resistance are not of critical significance in a digital switch, these parameters are of immense importance in *analog switches*. The use of MOSFETs to implement analog switches will be discussed in Chapter 10. Also, we shall shortly see that in practical circuit implementations of the DAC, the binary-weighted currents are generated by current sources. In this case the analog switch can be realized using the differential-pair circuit, as will be shown shortly.

A disadvantage of the binary-weighted resistor network is that for a large number of bits ($N > 4$) the spread between the smallest and largest resistances becomes quite large. This implies difficulties in maintaining accuracy in resistor values. A more convenient scheme exists utilizing a resistive network called the *R*-2*R* ladder.

9.8.2 *R*-2*R* Ladders

Figure 9.40 shows the basic arrangement of a DAC using an *R*-2*R* ladder. Because of the small spread in resistance values, this network is usually preferred to the binary-weighted scheme discussed earlier, especially for $N > 4$. Operation of the *R*-2*R* ladder is straightforward. First, it can be shown, by starting from the right and working toward the left, that the

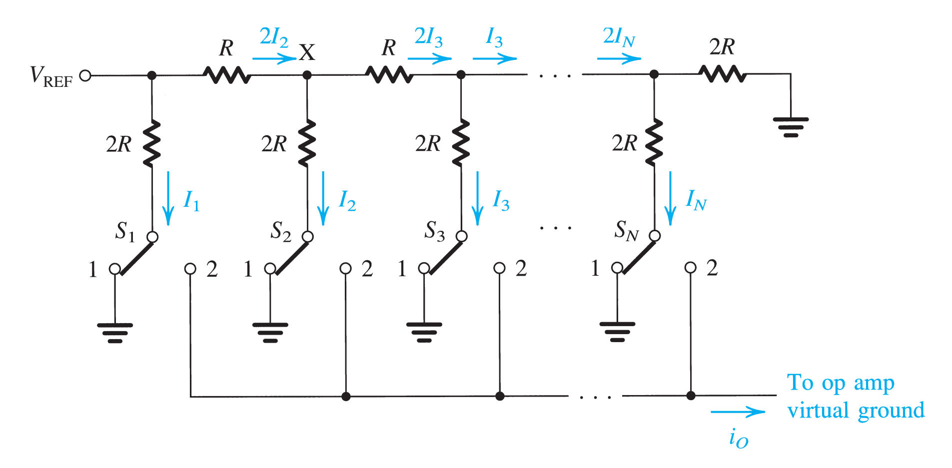


FIGURE 9.40 The basic circuit configuration of a DAC utilizing an *R*-2*R* ladder network.

resistance to the right of each ladder node, such as that labeled X, is equal to $2R$. Thus the current flowing to the right, away from each node, is equal to the current flowing downward to ground, and twice that current flows into the node from the left side. It follows that

$$I_1 = 2I_2 = 4I_3 = \cdots = 2^{N-1} I_N \tag{9.112}$$

Thus, as in the binary-weighted resistive network, the currents controlled by the switches are binary weighted. The output current i_O will therefore be given by

$$i_O = \frac{V_{REF}}{R} D \tag{9.113}$$

9.8.3 A Practical Circuit Implementation

A practical circuit implementation of the DAC utilizing an R-$2R$ ladder is shown in Fig. 9.41. The circuit utilizes BJTs to generate binary-weighted constant currents $I_1, I_2, \ldots, I_N$, which are switched between ground and virtual ground of an output summing op amp (not shown). We shall first show that the currents I_1 to I_N are indeed binary-weighted, with I_1 corresponding to the MSB and I_N corresponding to the LSB of the DAC.

Starting at the two rightmost transistors, Q_N and Q_t, we see that if they are matched, their emitter currents will be equal and are denoted (I_N/α). Transistor Q_t is included to provide proper termination of the R-$2R$ network. The voltage between the base line of the BJTs and node N will be

$$V_N = V_{BE_N} + \left(\frac{I_N}{\alpha}\right)(2R)$$

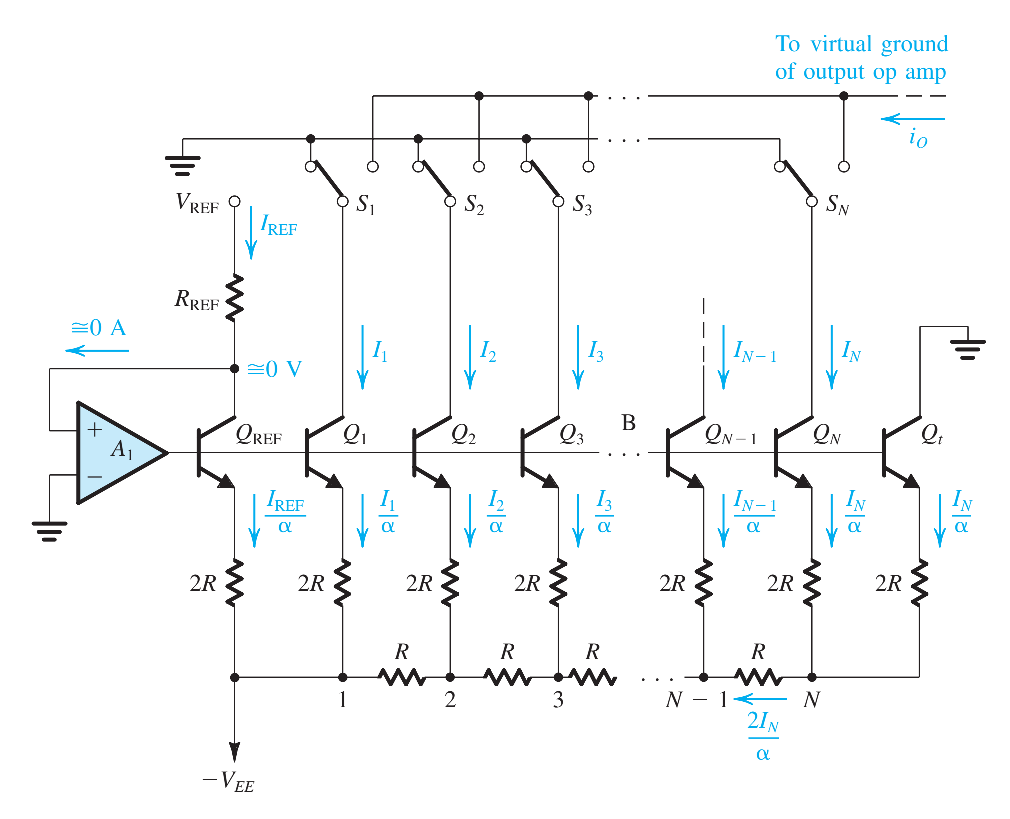


FIGURE 9.41 A practical circuit implementation of a DAC utilizing an R-$2R$ ladder network.

where V_{BE_N} is the base–emitter voltage of Q_N. Since the current flowing through the resistor R connected to node N is $(2I_N/\alpha)$, the voltage between node B and node $(N-1)$ will be

$$V_{N-1} = V_N + \left(\frac{2I_N}{\alpha}\right)R = V_{BE_N} + \frac{4I_N}{\alpha}R$$

Assuming, for the moment, that $V_{BE_{N-1}} = V_{BE_N}$, we see that a voltage of $(4I_N/\alpha)R$ appears across the resistance $2R$ in the emitter of Q_{N-1}. Thus Q_{N-1} will have an emitter current of $(2I_N/\alpha)$ and a collector current of $(2I_N)$, twice the current in Q_N. The two transistors will have equal V_{BE} drops if their junction areas are scaled in the same proportion as their currents, which is usually done in practice.

Proceeding in the manner above we can show that

$$I_1 = 2I_2 = 4I_3 = \cdots = 2^{N-1}I_N \tag{9.114}$$

under the assumption that the EBJ areas of Q_1 to Q_N are scaled in a binary-weighted fashion.

Next consider op amp A_1, which, together with the reference transistor Q_{REF}, forms a negative-feedback loop. (Convince yourself that the feedback is indeed negative.) A virtual ground appears at the collector of Q_{REF} forcing it to conduct a collector current $I_{\text{REF}} = V_{\text{REF}}/R_{\text{REF}}$ independent of whatever imperfections Q_{REF} might have. Now, if Q_{REF} and Q_1 are matched, their collector currents will be equal,

$$I_1 = I_{\text{REF}}$$

Thus, the binary-weighted currents are directly related to the reference current, independent of the exact values of V_{BE} and α. Also observe that op amp A_1 supplies the base currents of all the BJTs.

9.8.4 Current Switches

Each of the single-pole double-throw switches in the DAC circuit of Fig. 9.41 can be implemented by a circuit such as that shown in Fig. 9.42 for switch S_m. Here I_m denotes the current flowing in the collector of the mth-bit transistor. The circuit is a differential pair with the

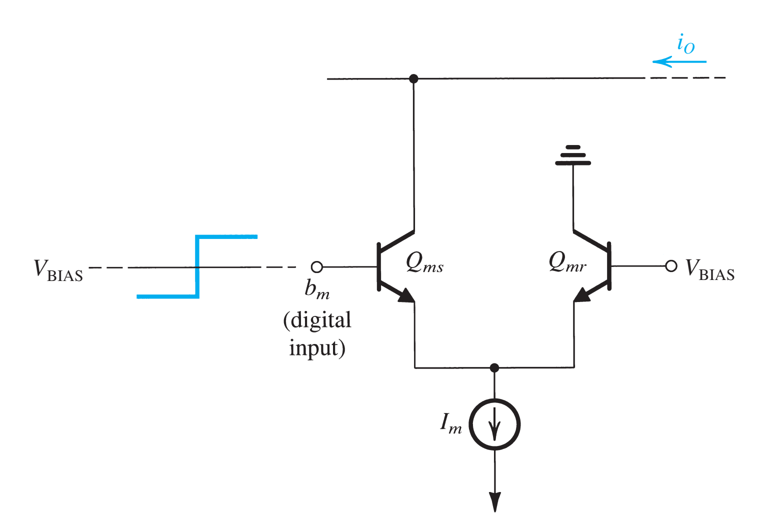


FIGURE 9.42 Circuit implementation of switch S_m in the DAC of Fig. 9.41. In a BiCMOS technology, Q_{ms} and Q_{mr} can be implemented using MOSFETs, thus avoiding the inaccuracy caused by the base current of BJTs.

base of the reference transistor Q_{mr} connected to a suitable dc voltage V_{BIAS}, and the digital signal representing the mth bit b_m applied to the base of the other transistor Q_{ms}. If the voltage representing b_m is higher than V_{BIAS} by a few hundred millivolts, Q_{ms} will turn on and Q_{mr} will turn off. The bit current I_m will flow through Q_{ms} and onto the output summing line. On the other hand, when b_m is low, Q_{ms} will be off and I_m will flow through Q_{mr} to ground.

The current switch of Fig. 9.42 is simple and features high-speed operation. It suffers, however, from the fact that part of the current I_m flows through the base of Q_{ms} and thus does not appear on the output summing line. More elaborate circuits for current switches can be found in Grebene (1984). Also, in a BiCMOS technology the differential-pair transistors Q_{ms} and Q_{mr} can be replaced with MOSFETs, thus eliminating the base current problem.

EXERCISES

9.32 What is the maximum resistor ratio required by a 12-bit D/A converter utilizing a binary-weighted resistor network?

Ans. 2048

9.33 If the input bias current of an op amp, used as the output summer in a 10-bit DAC, is to be no more than that equivalent to $\frac{1}{4}$ LSB, what is the maximum current required to flow in R_f for an op amp whose bias current is as great as 0.5 μA?

Ans. 2.046 mA



9.9 A/D CONVERTER CIRCUITS

There exist a number of A/D conversion techniques varying in complexity and speed. We shall discuss four different approaches: two simple, but slow, schemes, one complex (in terms of the amount of circuitry required) but extremely fast method, and, finally, a method particularly suited for MOS implementation.

9.9.1 The Feedback-Type Converter

Figure 9.43 shows a simple A/D converter that employs a comparator, an up/down counter, and a D/A converter. The comparator circuit provides an output that assumes one of two

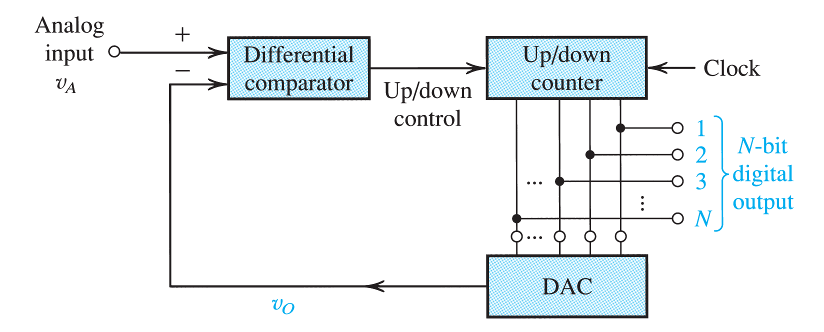


FIGURE 9.43 A simple feedback-type A/D converter.

distinct values: positive when the difference input signal is positive, and negative when the difference input signal is negative. We shall study comparator circuits in Chapter 13. An up/down counter is simply a counter that can count either up or down depending on the binary level applied at its up/down control terminal. Because the A/D converter of Fig. 9.43 employs a DAC in its feedback loop it is usually called a feedback-type A/D converter. It operates as follows: With a 0 count in the counter, the D/A converter output, v_O, will be zero and the output of the comparator will be high, instructing the counter to count the clock pulses in the up direction. As the count increases, the output of the DAC rises. The process continues until the DAC output reaches the value of the analog input signal, at which point the comparator switches and stops the counter. The counter output will then be the digital equivalent of the input analog voltage.

Operation of the converter of Fig. 9.43 is slow if it starts from zero. This converter however, tracks incremental changes in the input signal quite rapidly.

9.9.2 The Dual-Slope A/D Converter

A very popular high-resolution (12- to 14-bit) (but slow) A/D conversion scheme is illustrated in Fig. 9.44. To see how it operates, refer to Fig. 9.44 and assume that the analog input signal v_A is negative. Prior to the start of the conversion cycle, switch S_2 is closed, thus discharging capacitor C and setting $v_1 = 0$. The conversion cycle begins with opening S_2 and connecting the integrator input through switch S_1 to the analog input signal. Since v_A is negative, a current $I = v_A/R$ will flow through R in the direction away from the integrator. Thus v_1 rises linearly with a slope of $I/C = v_A/RC$, as indicated in Fig. 9.44(b). Simultaneously, the counter is enabled and it counts the pulses from a fixed-frequency clock. This phase of the conversion process continues for a fixed duration T_1. It ends when the counter has accumulated a fixed count denoted n_{REF}. Usually, for an N-bit converter, $n_{REF} = 2^N$. Denoting the peak voltage at the output of the integrator as V_{PEAK}, we can write with reference to Fig. 9.44(b)

$$\frac{V_{PEAK}}{T_1} = \frac{v_A}{RC} \tag{9.115}$$

At the end of this phase, the counter is reset to zero.

Phase II of the conversion begins at $t = T_1$ by connecting the integrator input through switch S_1 to the positive reference voltage V_{REF}. The current into the integrator reverses direction and is equal to V_{REF}/R. Thus v_1 decreases linearly with a slope of (V_{REF}/RC). Simultaneously the counter is enabled and it counts the pulses from the fixed-frequency clock. When v_1 reaches zero volts, the comparator signals the control logic to stop the counter. Denoting the duration of phase II by T_2, we can write, by reference to Fig. 9.44(b),

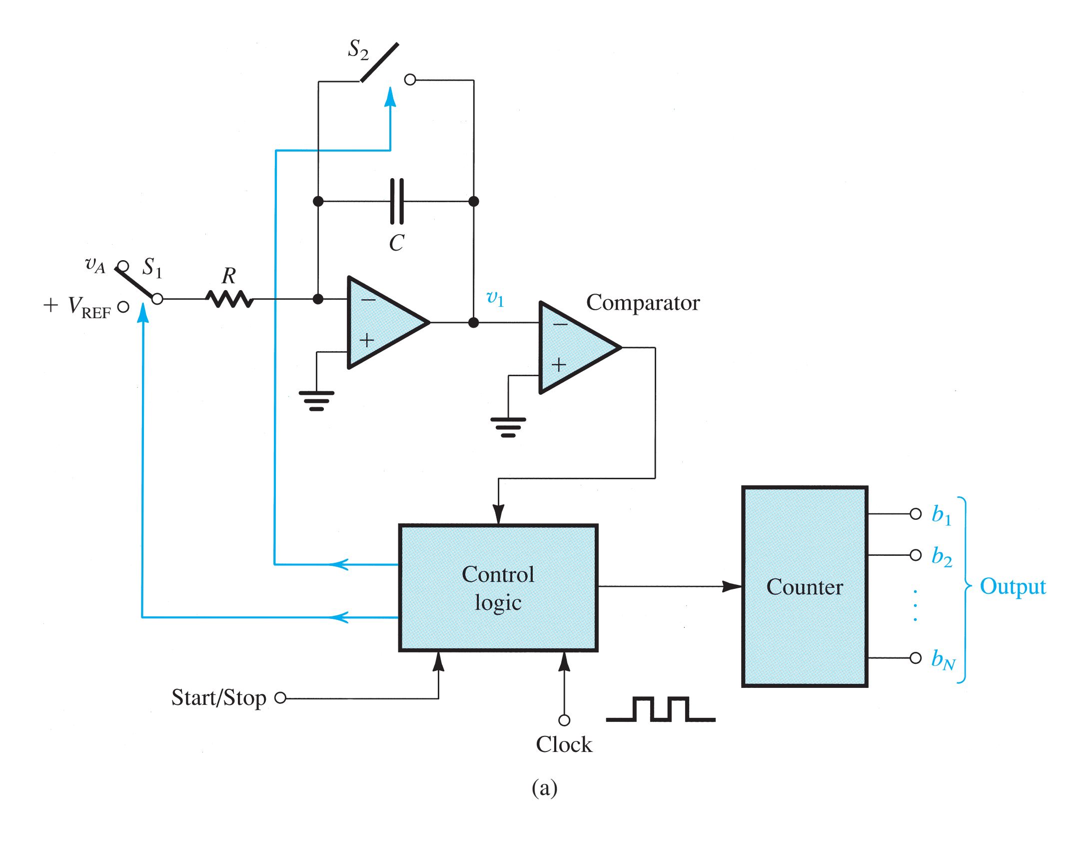
$$\frac{V_{PEAK}}{T_2} = \frac{V_{REF}}{RC} \tag{9.116}$$

Equations (9.115) and (9.116) can be combined to yield

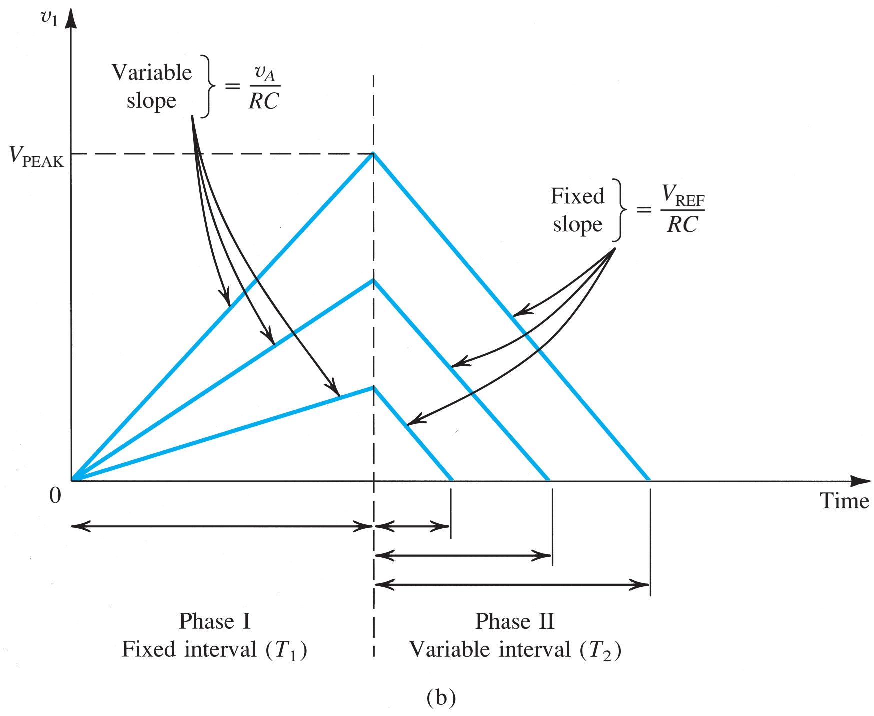
$$T_2 = T_1\left(\frac{v_A}{V_{REF}}\right) \tag{9.117}$$

Since the counter reading, n_{REF}, at the end of T_1 is proportional to T_1 and the reading, n, at the end of T_2 is proportional to T_2, we have

$$n = n_{REF}\left(\frac{v_A}{V_{REF}}\right) \tag{9.118}$$



(a)



(b)

FIGURE 9.44 The dual-slope A/D conversion method. Note that v_A is assumed to be negative.

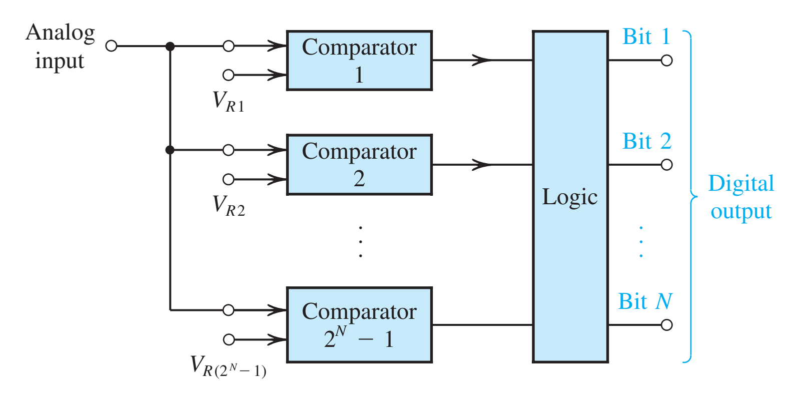


FIGURE 9.45 Parallel, simultaneous, or flash A/D conversion.

Thus the content of the counter,[5] n, at the end of the conversion process is the digital equivalent of v_A.

The dual-slope converter features high accuracy, since its performance is independent of the exact values of R and C. There exist many commercial implementations of the dual-slope method, some of which utilize CMOS technology.

9.9.3 The Parallel or Flash Converter

The fastest A/D conversion scheme is the simultaneous, parallel, or flash conversion process illustrated in Fig. 9.45. Conceptually, flash conversion is very simple. It utilizes $2^N - 1$ comparators to compare the input signal level with each of the $2^N - 1$ possible quantization levels. The outputs of the comparators are processed by an encoding-logic block to provide the N bits of the output digital word. Note that a complete conversion can be obtained within one clock cycle.

Although flash conversion is very fast, the price paid is a rather complex circuit implementation. Variations on the basic technique have been successfully employed in the design of IC converters.

9.9.4 The Charge-Redistribution Converter

The last A/D conversion technique that we shall discuss is particularly suited for CMOS implementation. As shown in Fig. 9.46, the circuit utilizes a binary-weighted capacitor array, a voltage comparator, and analog switches; control logic (not shown in Fig. 9.46) is also required. The circuit shown is for a 5-bit converter; capacitor C_T serves the purpose of terminating the capacitor array, making the total capacitance equal to the desired value of $2C$.

Operation of the converter can be divided into three distinct phases, as illustrated in Fig. 9.46. In the sample phase (Fig. 9.46a) switch S_B is closed, thus connecting the top plate of all capacitors to ground and setting v_O to zero. Meanwhile, switch S_A is connected to the analog input voltage v_A. Thus the voltage v_A appears across the total capacitance of $2C$, resulting in a stored charge of $2Cv_A$. Thus, during this phase, a sample of v_A is taken and a proportional amount of charge is stored on the capacitor array.

[5] Note that n is *not* a continuous function of v_A, as might be inferred from Eq. (9.118). Rather, n takes on discrete values corresponding to one of the 2^N quantized levels of v_A.

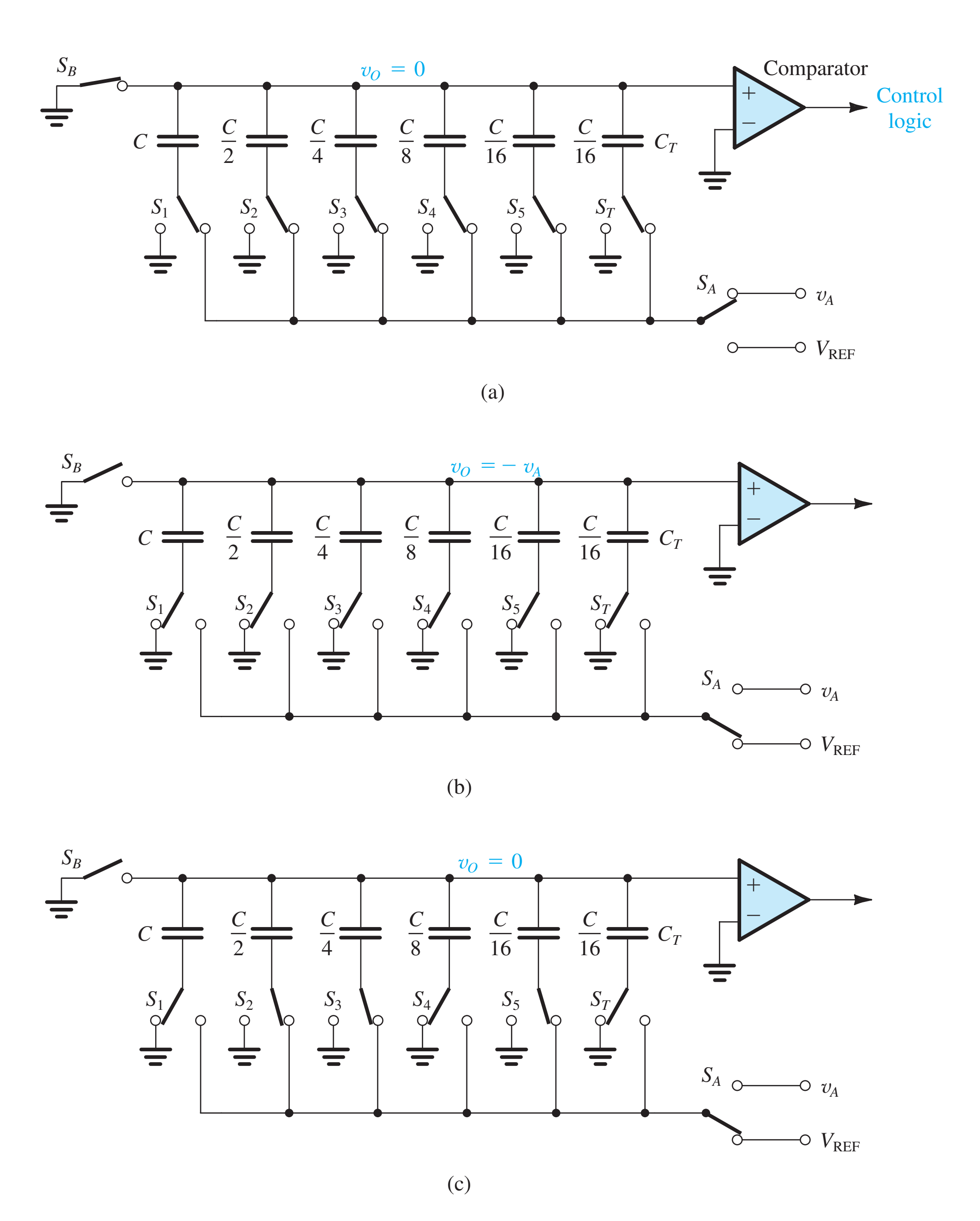


FIGURE 9.46 Charge-redistribution A/D converter suitable for CMOS implementation: **(a)** sample phase, **(b)** hold phase, and **(c)** charge-redistribution phase.

During the hold phase (Fig. 9.46b), switch S_B is opened and switches S_1 to S_5, and S_T are thrown to the ground side. Thus the top plate of the capacitor array is open-circuited while their bottom plates are connected to ground. Since no discharge path has been provided, the capacitor charges must remain constant, with the total equal to $2Cv_A$. It follows that the voltage at the top plate must become $-v_A$. Finally, note that during the hold phase, S_A is connected to V_{REF} in preparation for the charge-redistribution phase.

Next, we consider the operation during the charge-redistribution phase illustrated in Fig. 9.46(c). First, switch S_1 is connected to V_{REF} (through S_A). The circuit then consists of V_{REF}, a series capacitor C, and a total capacitance to ground of value C. This capacitive divider causes a voltage increment of $V_{REF}/2$ to appear on the top plates. Now, if v_A is greater than $V_{REF}/2$, the net voltage at the top plate will remain negative, which means that S_1 will be left in its new position as we move on to switch S_2. If, on the other hand, v_A was smaller than $V_{REF}/2$, then the net voltage at the top plate would become positive. The comparator will detect this situation and signal the control logic to return S_1 to its ground position and then to move on to S_2.

Next, switch S_2 is connected to V_{REF}, which causes a voltage increment of $V_{REF}/4$ to appear on the top plate. If the resulting voltage is still negative, S_2 is left in its new position; otherwise, S_2 is returned to its ground position. We then move on to switch S_3, and so on until all the bit switches S_1 to S_5 have been tried.

It can be seen that during the charge-redistribution phase the voltage on the top plate will be reduced incrementally to zero. The connection of the bit switches at the conclusion of this phase gives the output digital word; a switch connected to ground indicates a 0 value for the corresponding bit, whereas connection to V_{REF} indicates a 1. The particular switch configuration depicted in Fig. 9.46(c) is for $D = 01101$. Observe that at the end of the conversion process, all the charge is stored in the capacitors corresponding to 1 bits; the capacitors of the 0 bits have been discharged.

The accuracy of this A/D conversion method is independent of the value of stray capacitances from the bottom plate of the capacitors to ground. This is because the bottom plates are connected either to ground or to V_{REF}; thus the charge on the stray capacitances will not flow into the capacitor array. Also, because both the initial and the final voltages on the top plate are zero, the circuit is also insensitive to the stray capacitances between the top plates and ground.[6] The insensitivity to stray capacitances makes the charge-redistribution technique a reasonably accurate method capable of implementing A/D converters with as many as 10 bits.

EXERCISES

9.34 Consider the 5-bit charge-redistribution converter in Fig. 9.46 with $V_{REF} = 4$ V. What is the voltage increment appearing on the top plate when S_5 is switched? What is the full-scale voltage of this converter? If $v_A = 2.5$ V, which switches will be connected to V_{REF} at the end of conversion?

Ans. $\frac{1}{8}$ V; $\frac{31}{8}$ V; S_1 and S_3

9.35 Express the maximum quantization error of an N-bit A/D converter in terms of its least-significant bit (LSB) and in terms of its full-scale analog input V_{FS}.

Ans. $\pm\frac{1}{2}$ LSB; $V_{FS}/2(2^N - 1)$



9.10 SPICE SIMULATION EXAMPLE

We conclude this chapter with an example to illustrate the use of SPICE in the simulation of the two-stage CMOS op amp.

[6] More precisely, the final voltage can deviate from zero by as much as the analog equivalent of the LSB. Thus, the insensitivity to top-plate capacitance is not complete.

5.11 THE JUNCTION FIELD-EFFECT TRANSISTOR (JFET)

The **junction field-effect transistor,** or JFET, is perhaps the simplest transistor available. It has some important characteristics, notably a very high input resistance. Unfortunately, however (for the JFET), the MOSFET has an even higher input resistance. This, together with the many other advantages of MOS transistors, has made the JFET virtually obsolete. Currently, its applications are limited to discrete-circuit design, where it is used both as an amplifier and as a switch. Its integrated-circuit applications are limited to the design of the differential input stage of some operational amplifiers, where advantage is taken of its high input resistance (compared to the BJT). In this section, we briefly consider JFET operation and characteristics. Another important reason for including the JFET in the study of electronics is that it helps in understanding the operation of gallium arsenide devices, the subject of the next section.

Device Structure

As with other FET types, the JFET is available in two polarities: *n*-channel and *p*-channel. Fig. 5.69(a) shows a simplified structure of the *n*-channel JFET. It consists of a slab of *n*-type silicon with *p*-type regions diffused on its two sides. The *n* region is the channel, and the *p*-type regions are electrically connected together and form the gate. The device operation is based on reverse-biasing the *pn* junction between gate and channel. Indeed, it is the reverse bias on this junction that is used to control the channel width and hence the current flow from drain to source. The major role that this *pn* junction plays in the operation of this FET has given rise to its name: Junction Field-Effect Transistor (JFET).

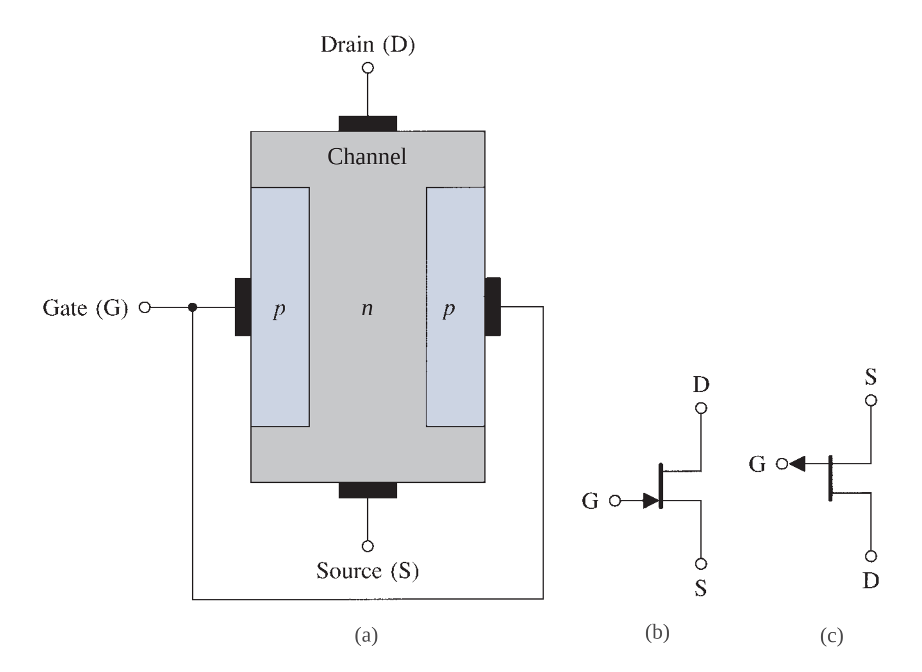


FIGURE 5.69 **(a)** Basic structure of *n*-channel JFET. This is a simplified structure utilized to explain device operation. **(b)** Circuit symbol for the *n*-channel JFET. **(c)** Circuit symbol for the *p*-channel JFET.

It should be obvious that a *p*-channel device can be fabricated by simply reversing all the semiconductor types, thus using *p*-type silicon for the channel and *n*-type silicon for the gate regions.

Figures 5.69(b) and (c) show the circuit symbols for JFETs of both polarities. Observe that the device polarity (*n*-channel or *p*-channel) is indicated by the direction of the arrowhead on the gate line. This arrowhead points in the forward direction of the gate–channel *pn* junction. Although the JFET is a symmetrical device whose source and drain can be interchanged, it is useful in circuit design to designate one of these two terminals as source and the other as drain. The circuit symbol achieves this designation by placing the gate closer to the source than to the drain.

Physical Operation

Consider an *n*-channel JFET and refer to Fig. 5.70(a). (Note that to simplify matters, we will not show the electrical connection between the gate terminals; it is assumed, however, that the two terminals labeled G are joined together.) With $v_{GS} = 0$, the application of a voltage v_{DS} causes current to flow from the drain to the source. When a negative v_{GS} is applied, the depletion region of the gate–channel junction widens and the channel becomes correspondingly narrower; thus the channel resistance increases and the current i_D (for a given v_{DS}) decreases. Because v_{DS} is small, the channel is almost of uniform width. The JFET is simply operating as a resistance whose value is controlled by v_{GS}. If we keep increasing v_{GS} in the negative direction, a value is reached at which the depletion region occupies the entire channel. At this value of v_{GS} the channel is completely depleted of charge carriers (electrons); the channel has in effect disappeared. This value of v_{GS} is therefore the threshold voltage of the device, V_t, which is obviously negative for an *n*-channel JFET. For JFETs the threshold voltage is called the **pinch-off voltage** and is denoted V_P.

Consider next the situation depicted in Fig. 5.70(b). Here v_{GS} is held constant at a value greater (that is, less negative) than V_P, and v_{DS} is increased. Since v_{DS} appears as a voltage drop across the length of the channel, the voltage increases as we move along the channel from source to drain. It follows that the reverse-bias voltage between gate and channel varies at different points along the channel and is highest at the drain end. Thus the channel acquires a tapered shape and the i_D–v_{DS} characteristic becomes nonlinear. When the reverse bias at the drain end, v_{GD}, falls below the pinch-off voltage V_P, the channel is pinched off at the drain end and the drain current saturates. The remainder of the description of JFET operation follows closely that given for the depletion MOSFET.

The description above clearly indicates that the JFET is a depletion-type device. Its characteristics should therefore be similar to those of the depletion-type MOSFET. This is true with a very important exception: While it is possible to operate the depletion-type MOSFET in the enhancement mode (by simply applying a positive v_{GS} if the device is *n* channel) this is impossible in the JFET case. If we attempt to apply a positive v_{GS}, the gate–channel *pn* junction becomes forward biased and the gate ceases to control the channel. Thus the maximum v_{GS} is limited to 0 V, though it is possible to go as high as 0.3 V or so since a *pn* junction remains essentially cut off at such a small forward voltage.

Current–Voltage Characteristics

The current–voltage characteristics of the JFET are identical to those of the depletion-mode MOSFET studied in Section 5.3 except that for the JFET the maximum v_{GS} allowed is

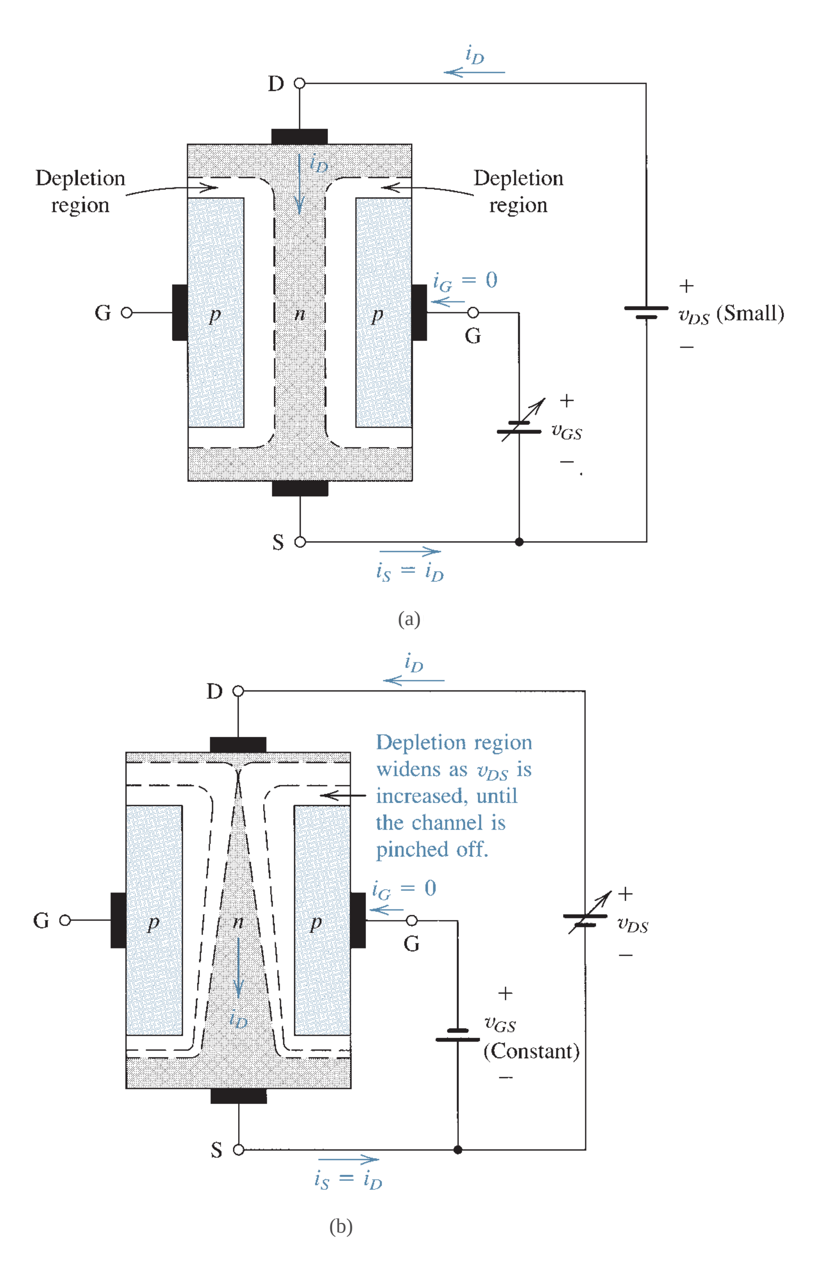


FIGURE 5.70 Physical operation of the *n*-channel JFET: **(a)** For small v_{DS} the channel is uniform and the device functions as a resistance whose value is controlled by v_{GS}. **(b)** Increasing v_{DS} causes the channel to acquire a tapered shape and eventually pinch-off occurs. Note that, though not shown, the two gate regions are electrically connected.

normally 0 V. Furthermore, the JFET is specified in terms of the pinch-off voltage V_P (equal to V_t of the MOSFET) and the *d*rain-to-*s*ource current with the gate *s*horted to the source, I_{DSS}, which corresponds to $\frac{1}{2}k_n'V_t^2$ for the MOSFET. With these substitutions, the *n*-channel JFET characteristics can be described as follows:

Cutoff: $v_{GS} \leq V_P$, $i_D = 0$

Triode region: $V_P \le v_{GS} \le 0$, $v_{DS} \le v_{GS} - V_P$

$$i_D = I_{DSS}\left[2\left(1-\frac{v_{GS}}{V_P}\right)\left(\frac{v_{DS}}{-V_P}\right)-\left(\frac{v_{DS}}{V_P}\right)^2\right] \tag{5.116}$$

Saturation (pinch-off) region: $V_P \le v_{GS} \le 0$, $v_{DS} \ge v_{GS} - V_P$

$$i_D = I_{DSS}\left(1-\frac{v_{GS}}{V_P}\right)^2(1+\lambda v_{DS}) \tag{5.117}$$

where λ is the inverse of the Early voltage; $\lambda = 1/V_A$, and V_A and λ are positive for *n*-channel devices.

Recalling that for an *n*-channel device, V_P is negative, we see that operation in the pinch-off region is obtained when the drain voltage is greater than the gate voltage by at least $|V_P|$.

Since the gate–channel junction is always reverse-biased, only a leakage current flows through the gate terminal. From Chapter 3, we know that such a current is of the order of 10^{-9} A. Although i_G is very small, and is assumed zero in almost all applications, it should be noted that the gate current in a JFET is many orders of magnitude greater than the gate current in a MOSFET. Of course the latter is so tiny because of the insulated gate structure. Another complication arises in the JFET because of the strong dependence of gate leakage current on temperature—approximately doubling for every 10°C rise in temperature, just as in the case of a reverse-biased diode (see Chapter 3).

The *p*-Channel JFET

The current–voltage characteristics of the *p*-channel JFET are described by the same equations as the *n*-channel JFET. Note, however, that for the *p*-channel JFET, V_P is positive, $0 \le v_{GS} \le V_P$, v_{DS} is negative, λ and V_A are negative, and the current i_D flows out of the drain terminal. To operate the *p*-channel JFET in pinch-off, $v_{DS} \le v_{GS} - V_P$, which in words means that the drain voltage must be lower than the gate voltage by at least $|V_P|$. Otherwise, with $v_{DS} \ge v_{GS} - V_P$, the *p*-channel JFET operates in the triode region.

The JFET Small-Signal Model

The JFET small-signal model is identical to that of the MOSFET [see Fig. 5.34(b)]. Here, g_m is given by

$$g_m = \left(\frac{2I_{DSS}}{|V_P|}\right)\left(1-\frac{V_{GS}}{V_P}\right) \tag{5.118a}$$

or alternatively by

$$g_m = \left(\frac{2I_{DSS}}{|V_P|}\right)\sqrt{\frac{I_D}{I_{DSS}}} \tag{5.118b}$$

where V_{GS} and I_D are the dc bias quantities, and

$$r_o = \frac{|V_A|}{I_D} \tag{5.119}$$

At high frequencies, the equivalent circuit of Fig. 5.67(c) applies with C_{gs} and C_{gd} being both depletion capacitances. Typically, $C_{gs} = 1$ to 3 pF, $C_{gd} = 0.1$ to 0.5 pF, and $f_T = 20$ to 100 MHz.

EXERCISES

In Exercises 5.43 to 5.46, let the *n*-channel JFET have $V_P = -4$ V and $I_{DSS} = 10$ mA, and unless otherwise specified assume that in pinch-off (saturation) the output resistance is infinite.

5.43 For $v_{GS} = -2$ V, find the minimum v_{DS} for the device to operate in pinch-off. Calculate i_D for $v_{GS} = -2$ V and $v_{DS} = 3$ V.

Ans. 2 V; 2.5 mA

5.44 For $v_{DS} = 3$ V, find the change in i_D corresponding to a change in v_{GS} from –2 to –1.6 V.

Ans. 1.1 mA

5.45 For small v_{DS}, calculate the value of r_{DS} at $v_{GS} = 0$ V and at $v_{GS} = -3$ V.

Ans. 200 Ω; 800 Ω

5.46 If $V_A = 100$ V, find the JFET output resistance r_o when operating in pinch-off at a current of 1 mA, 2.5 mA, and 10 mA.

Ans. 100 kΩ; 40 kΩ; 10 kΩ

D5.47 The JFET in the circuit of Fig. E5.47 has $V_P = -3$ V, $I_{DSS} = 9$ mA, and $\lambda = 0$. Find the values of all resistors so that $V_G = 5$ V, $I_D = 4$ mA, and $V_D = 11$ V. Design for 0.05 mA in the voltage divider.

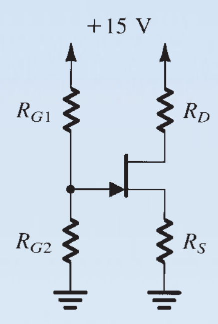


FIGURE E5.47

Ans. $R_{G1} = 200$ kΩ; $R_{G2} = 100$ kΩ; $R_S = 1.5$ kΩ; $R_D = 1$ kΩ

5.48 For the JFET circuit designed in Exercise 5.47, let an input signal v_i be capacitively coupled to the gate, a large bypass capacitor be connected between the source and ground, and the output signal v_o be taken from the drain through a large coupling capacitor. The resulting common-source amplifier is shown in Fig. E5.48. Calculate g_m and r_o (assuming $V_A = 100$ V). Also find R_i, $A_v \equiv (v_o/v_i)$, and R_o.

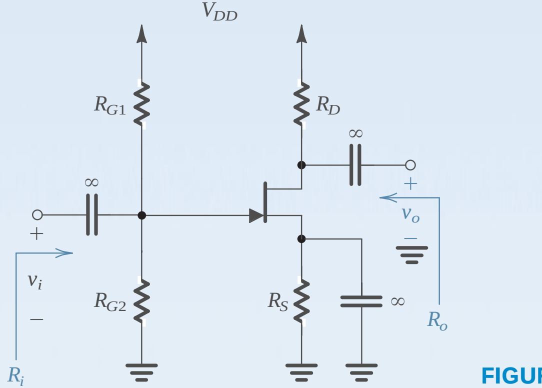


FIGURE E5.48

Ans. 4 mA/V; 25 kΩ; 66.7 kΩ; –3.8 V/V; 962 Ω



5.12 GALLIUM ARSENIDE (GaAs) DEVICES—THE MESFET[10]

The devices discussed thus far, and indeed the devices used in most of the circuits studied in this book, are made of silicon. This reflects the situation that has existed in the microelectronics industry for at least three decades. Furthermore, owing to the advances that are continually being made in silicon device and circuit technologies, the dominance of silicon as the most useful semiconductor material is expected to continue for many years to come. Nevertheless, another semiconductor material has been making inroads into digital applications that require extremely high speeds of operation and analog applications that require very high operating frequencies. We refer to gallium arsenide (GaAs), a compound semiconductor formed of gallium, which is in the third column of the periodic table of elements, and arsenic, which is in the fifth column; thus GaAs is known as a III-V semiconductor.

The major advantage that GaAs offers over silicon is that electrons travel much faster in *n*-type GaAs than in silicon. This is a result of the fact that the electron drift mobility μ_n (which is the constant that relates the electron drift velocity to the electric field; velocity $= \mu_n E$) is five to ten times higher in GaAs than in silicon. Thus for the same input voltages, GaAs devices have higher output currents, and thus higher g_m, than the corresponding silicon devices. The larger output currents enable faster charging and discharging of load and parasitic capacitances and thus result in increased speeds of operation.

Gallium arsenide devices have been used for some years in the design of discrete-component amplifiers for microwave applications (in the 10^9 Hz or GHz frequency range). More recently, GaAs has begun to be employed in the design of very-high-speed digital integrated circuits and in analog ICs, such as op amps, that operate in the hundreds of MHz frequency range. Although the technology is still relatively immature, suffering from yield and reliability problems and generally limited to low levels of integration, it offers great potential. Therefore, this book includes a brief study of GaAs devices and circuits. Specifically, the basic GaAs devices are studied in this section; their basic amplifier circuit configurations are discussed in Section 6.8; and GaAs digital circuits are studied in Section 14.8.

The Basic GaAs Devices

Although there are a number of GaAs technologies currently in various stages of development, we shall study the most mature of these technologies. The active device available in this technology is an *n*-channel field effect transistor known as the **metal semiconductor FET** or **MESFET.** The technology also provides a type of diode known as the **Schottky-barrier diode (SBD).** (Recall that the SBD was briefly introduced in Section 3.9.) The structure of these two basic devices is illustrated by their cross sections, depicted in Fig. 5.71. The GaAs circuit is formed on an undoped GaAs substrate. Since the conductivity of undoped GaAs is very low, the substrate is said to be semi-insulating. This turns out to be an advantage for GaAs technology as it simplifies the process of isolating the devices on the chip from one another, as well as resulting in smaller parasitic capacitances between the devices and the circuit ground.

As indicated in Fig. 5.71, a Schottky-barrier diode consists of a metal–semiconductor junction. The metal, referred to as the Schottky-barrier metal to distinguish it from the different kind of metal used to make a contact (see Long and Butner (1990) for a detailed explanation of the difference), forms the anode of the diode. The *n*-type GaAs forms the

[10] The material in this section is required only for the study of the GaAs circuits in Sections 6.8 and 14.8. Otherwise, this section can be skipped without loss of continuity.

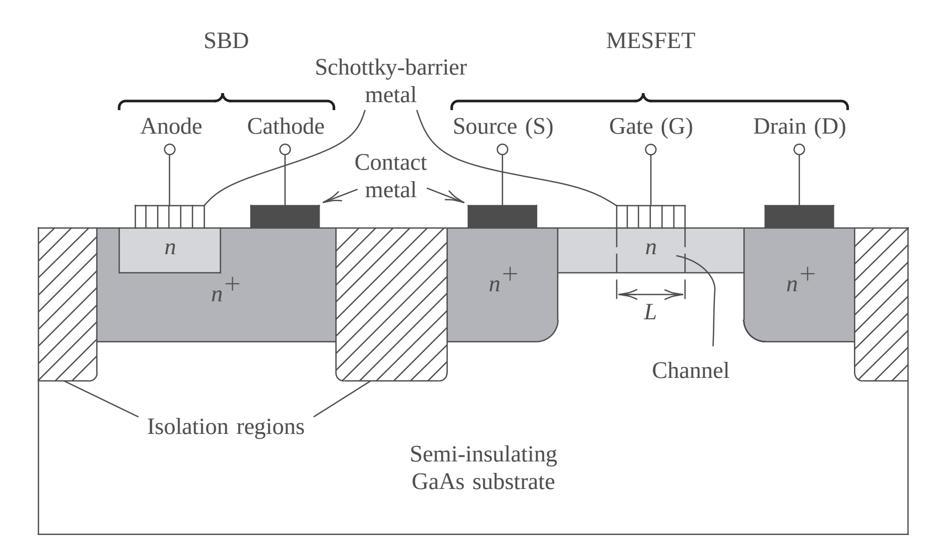


FIGURE 5.71 Cross-section of a GaAs Schottky-barrier diode (SBD) and a MESFET.

cathode. Note that heavily doped *n*-type GaAs (indicated by n^+) is used between the *n* region and the cathode metal contact in order to keep the parasitic series resistance low.

The gate of the MESFET is formed by Schottky-barrier metal in direct contact with the *n*-type GaAs that forms the channel region. The channel length *L* is defined by the length of the gate electrode, and similarly for the width *W* (in the direction perpendicular to the page). To reduce the parasitic resistances between the drain and source contacts and the channel, the two contacts are surrounded with heavily doped (n^+) GaAs.

Since the main reason for using GaAs circuits is to achieve high speed/frequency of operation, the channel length is made as small as possible. Typically $L = 0.2$–$2\ \mu$m. Also, usually all the transistors on the IC chip are made to have the same length, leaving only the width *W* of each device to be specified by the circuit designer.

Only *n*-channel MESFETs are available in GaAs technology. This is because holes have a relatively low drift mobility in GaAs, making *p*-channel MESFETs unattractive. The lack of complementary transistors is a definite disadvantage of GaAs technology. Correspondingly, it makes the task of the circuit designer even more challenging than usual.

Device Operation

The MESFET operates in a very similar manner to the JFET, with the Schottky metal playing the role of the *p*-type gate of the JFET (refer to Fig. 5.69). Basically, a depletion region forms in the channel below the gate surface, and the thickness of the depletion region is controlled by the gate voltage v_{GS}. This in turn effects control over the channel dimensions and thus on the current that flows from drain to source in response to an applied v_{DS}. The latter voltage causes the channel to have a tapered shape, with pinch-off eventually occurring at the drain end of the channel.

The most common GaAs MESFETs available are of the depletion type with a threshold voltage V_t (or, equivalently, pinch-off voltage V_P) in the range of –0.5 to –2.5 V. These devices can be operated with v_{GS} values ranging from the negative V_t to positive values as high as a few tenths of a volt. However, as v_{GS} reaches 0.7 V or so, the Schottky-barrier diode between gate and channel conducts heavily and the gate voltage no longer effectively

controls the drain-to-source current. Gate conduction, which is not possible in MOSFETs, is another definite disadvantage of the MESFET.

Although less common, enhancement-mode MESFETs are available in certain technologies. These normally-off devices are obtained by arranging that the depletion region existing at $v_{GS} = 0$ extends through the entire channel depth, thus blocking the channel and causing $i_D = 0$. To cause current to flow from drain to source the channel must be opened by applying to the gate a positive voltage of sufficient magnitude to reduce the thickness of the depletion region below that of the channel region. Typically, the threshold voltage V_t is between 0.1 and 0.3 V.

The above description of MESFET operation suggests that the i_D–v_{DS} characteristics should saturate at $v_{DS} = v_{GS} - V_t$, as is the case in a silicon JFET. It has been observed, however, that the i_D–v_{DS} characteristics of GaAs MESFETs saturate at lower values of v_{DS} and, furthermore, that the saturation voltages $v_{DS\text{sat}}$ do not depend strongly on the value of v_{GS}. This "early saturation" phenomenon comes about because the velocity of the electrons in the channel does not remain proportional to the electric field (which in turn is determined by v_{DS} and L; $E = v_{DS}/L$) as is the case in silicon; rather, the electron velocity reaches a high peak value and then saturates (that is, becomes constant independent of v_{DS}). The velocity-saturation effect is even more pronounced in short-channel devices ($L \le 1\ \mu$m), occurring at values of v_{DS} lower than $(v_{GS} - V_t)$.

Finally, a few words about the operation of the Schottky-barrier diode. Forward current is conducted by the majority carriers (electrons) flowing into the Schottky-barrier metal (the anode). Unlike the *pn*-junction diode, minority carriers play no role in the operation of the SBD. As a result, the SBD does not exhibit minority-carrier storage effects, which give rise to the diffusion capacitance of the *pn*-junction diode. Thus, the SBD has only one capacitive effect, that associated with the depletion-layer capacitance C_j.

Device Characteristics and Models

A first-order model for the MESFET, suitable for hand calculations, is obtained by neglecting the velocity-saturation effect, and thus the resulting model is almost identical to that of the JFET though expressed somewhat differently in order to correspond to the literature:

$$i_D = 0 \qquad \text{for } v_{GS} < V_t$$

$$i_D = \beta[2(v_{GS} - V_t)v_{DS} - v_{DS}^2](1 + \lambda v_{DS}) \qquad \text{for } v_{GS} \ge V_t,\ v_{DS} < v_{GS} - V_t$$

$$i_D = \beta(v_{GS} - V_t)^2(1 + \lambda v_{DS}) \qquad \text{for } v_{GS} \ge V_t,\ v_{DS} \ge v_{GS} - V_t \tag{5.120}$$

The only differences between these equations and those for the JFETs are (1) the channel-length modulation factor, $1 + \lambda v_{DS}$, is included also in the equation describing the triode region (also called the ohmic region) simply because λ of the MESFET is rather large and including this factor results in a better fit to measured characteristics; and (2) a transconductance parameter β is used so as to correspond with the MESFET literature. Obviously, β is related to I_{DSS} of the JFET and $k'(W/L)$ of the MOSFET. (Note, however, that this β has absolutely nothing to do with β of the BJT!)

A modification of this model to account for the early saturation effects is given in Hodges and Jackson (1988).

Figure 5.72(a) shows the circuit symbol for the depletion-type *n*-channel GaAs MESFET. Since only one type of transistor (*n* channel) is available, all devices will be drawn the

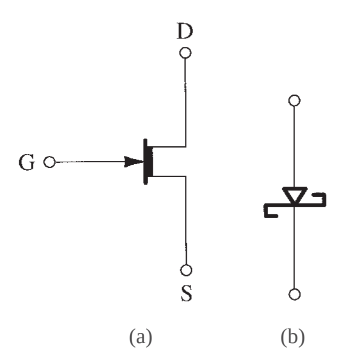


FIGURE 5.72 Circuit symbols for **(a)** an *n*-channel depletion-type GaAs MESFET, and **(b)** a Schottky-barrier diode (SBD).

same way, and there should be no confusion as to which terminal is the drain and which is the source.

The circuit symbol of the Schottky-barrier diode is depicted in Fig. 5.72(b). In spite of the fact that the physical operation of the SBD differs from that of the *pn*-junction diode, their *i*–*v* characteristics are identical. Thus the *i*–*v* characteristic of the SBD is given by the same exponential relationship studied in Chapter 3. For the GaAs SBD, the constant n is typically in the range of 1 to 1.2.

The small-signal model of the MESFET is identical to that of other FET types. The parameter values are given by

$$g_m = 2\beta(V_{GS} - V_t)(1 + \lambda V_{DS}) \tag{5.121}$$

$$r_o \equiv \left[\frac{\partial i_D}{\partial v_{DS}}\right]^{-1}$$

$$= 1/\lambda\beta(V_{GS} - V_t)^2 \tag{5.122}$$

The MESFET, however, has a rather high value for λ (0.1 to 0.3 V^{-1}) which results in a small output resistance r_o. This turns out to be a serious drawback of GaAs MESFET technology, resulting in low voltage-gain obtainable from each stage. Furthermore, it has been found that r_o decreases at high frequencies. Circuit design techniques for coping with the low r_o will be presented in Section 6.8.

For easy reference, Table 5.2 gives typical values for device parameters in a GaAs MESFET technology. The devices in this technology have a channel length $L = 1\ \mu$m. The values given are for a device with a width $W = 1\ \mu$m. The parameter values for actual devices can be obtained by appropriately scaling by the width W. This process is illustrated in the following example. Unless otherwise specified, the values of Table 5.2 are to be used for the exercises and the end-of-chapter problems.

TABLE 5.2 Typical Parameter Values for GaAs MESFETS and Schottky Diodes in $L = 1\ \mu$m Technology, Normalized for $W = 1\ \mu$m

$V_t = -1.0$ V
$\beta = 10^{-4}$ A/V^2
$\lambda = 0.1$ V^{-1}
$I_S = 10^{-15}$ A
$n = 1.1$

EXAMPLE 5.11

Figure 5.73 shows a simple GaAs MESFET amplifier, with the W values of the transistors indicated. Assume that the dc component of v_I, that is V_{GS1}, biases Q_1 at the current provided by the current source Q_2 so that both devices operate in saturation and that the dc output is at half of the supply voltage. Find:

(a) the β values for Q_1 and Q_2;

(b) V_{GS1};

(c) g_{m1}, r_{o1}, and r_{o2}; and

(d) the small-signal voltage gain.

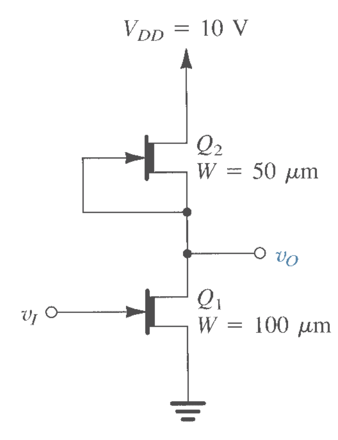


FIGURE 5.73 Circuit for Example 5.11: a simple MESFET amplifier.

Solution

(a) The values of β can be obtained by scaling the value given in Table 5.2 using the specified values of W,

$$\beta_1 = 100 \times 10^{-4} = 10^{-2} \text{ A/V}^2 = 10 \text{ mA/V}^2$$

$$\beta_2 = 50 \times 10^{-4} = 5 \times 10^{-3} \text{ A/V}^2 = 5 \text{ mA/V}^2$$

(b)

$$I_{D2} = \beta_2 (V_{GS2} - V_t)^2 (1 + \lambda V_{DS2})$$
$$= 5(0+1)^2(1+0.1\times 5)$$
$$= 7.5 \text{ mA}$$
$$I_{D1} = I_{D2} = 7.5 \text{ mA}$$
$$7.5 = \beta_1 (V_{GS1} - V_t)^2 (1 + \lambda V_{DS1})$$
$$= 10(V_{GS1} + 1)^2(1+0.1\times 5)$$

Thus,

$$V_{GS1} = -0.3 \text{ V}$$

(c)

$$g_{m1} = 2 \times 10(-0.3 + 1)(1 + 0.1 \times 5) = 21 \text{ mA/V}$$

$$r_{o1} = \frac{1}{0.1 \times 10(-0.3 + 1)^2} = 2 \text{ k}\Omega$$

$$r_{o2} = \frac{1}{0.1 \times 5(0 + 1)^2} = 2 \text{ k}\Omega$$

(d)

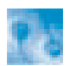
$$A_v = -g_{m1}(r_{o1} \,//\, r_{o2}) = -21 \times (2 \,//\, 2) = -21 \text{ V/V}$$

EXERCISE

5.49 For a MESFET with the gate shorted to the source and having $W = 10\ \mu$m, find the minimum voltage between drain and source to operate in saturation. For $V_{DS} = 5$ V, find the current I_D. What is the output resistance of this current source?

Ans. 1 V; 1.5 mA; 10 kΩ

As already mentioned, the main reason for using GaAs devices and circuits is their high frequency and high speed of operation. A remark is therefore in order on the internal capacitances and f_T of GaAs transistors. For a particular GaAs technology with $L = 1\ \mu$m, C_{gs} (at $V_{GS} = 0$ V) is 1.6 fF/μm-width, and C_{gd} (at $V_{DS} = 2$ V) is 0.16 fF/μm-width. Thus for a MESFET with $W = 100\ \mu$m, $C_{gs} = 0.16$ pF and $C_{gd} = 0.016$ pF. f_T typically ranges from 5 to 15 GHz.



6.8 GaAs AMPLIFIERS[3]

Gallium arsenide (GaAs) technology makes possible the design of amplifiers having very wide bandwidths, in the hundreds of megahertz or even gigahertz range. In this section we shall study some of the circuit design techniques that have been developed over the last few years for the design of GaAs amplifiers. As will be seen, these techniques aim to circumvent the major problem of the MESFET, namely, its low output resistance in saturation. Before proceeding with this section the reader is advised to review the material on GaAs devices presented in Section 5.12.

Current Sources

Current sources play a fundamental role in the design of integrated-circuit amplifiers, being employed both for biasing and as active loads. In GaAs technology, the simplest way to implement a current source is to connect the gate of a depletion-type MESFET to its source,

[3] This section can be omitted with no loss in continuity.

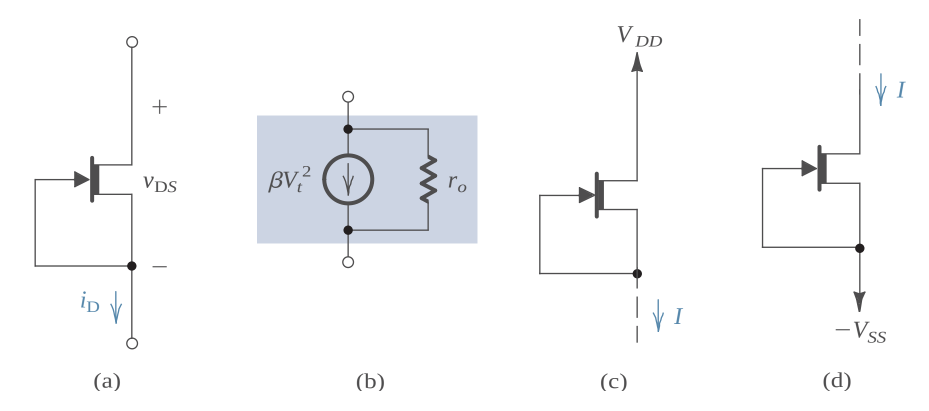


FIGURE 6.39 **(a)** The basic MESFET current source; **(b)** equivalent circuit of the current source; **(c)** the current source connected to a positive power supply to source currents to loads at voltages $\leq V_{DD} - |V_t|$; **(d)** the current source connected to a negative power supply to sink currents from loads at voltages $\geq -V_{SS} + |V_t|$.

as shown in Fig. 6.39(a). Provided that v_{DS} is maintained greater than $|V_t|$, the MESFET will operate in saturation and the current i_D will be

$$i_D = \beta V_t^2(1 + \lambda v_{DS}) \tag{6.126}$$

Thus the current source will have the equivalent circuit shown in Fig. 6.39(b), where the output resistance is the MESFET r_o,

$$r_o = 1/\lambda\beta V_t^2 \tag{6.127}$$

In JFET terminology, $\beta V_t^2 = I_{DSS}$ and $\lambda = 1/|V_A|$; thus

$$r_o = |V_A|/I_{DSS} \tag{6.128}$$

Since for the MESFET, λ is relatively high (0.1 to 0.3 V^{-1}) the output resistance of the current source of Fig. 6.39(a) is usually low, rendering this current-source realization inadequate for most applications. Before considering means for increasing the effective output resistance of the current source, we show in Fig. 6.39(c) how the basic current source can be connected to *source* currents to a load whose voltage can be as high as $V_{DD} - |V_t|$. Alternatively, the same device can be connected as shown in Fig. 6.39(d) to *sink* currents from a load whose voltage can be as low as $-V_{SS} + |V_t|$.

EXERCISE

6.23 Using the device data given in Table 5.2 (page 456), find the current provided by a 10-μm-wide MESFET connected in the current-source configuration. Let the source be connected to a −5-V supply and find the current when the drain voltage is –4V. What is the output resistance of the current source? What change in current occurs if the drain voltage is raised by +4V?

Ans. 1.1 mA; 10 kΩ; 0.4 mA

A Cascode Current Source

The output resistance of the current source can be increased by utilizing the cascode configuration as shown in Fig. 6.40. The output resistance R_o of the cascode current source can be

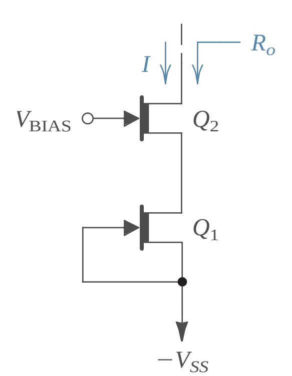


FIGURE 6.40 Adding the cascode transistor Q_2 increases the output resistance of the current source by the factor $g_{m2}r_{o2}$; that is, $R_o = g_{m2}r_{o2}r_{o1}$.

found by using Eq. (6.116),

$$R_o \simeq g_{m2}r_{o2}r_{o1} \tag{6.129}$$

Thus, adding the cascode transistor Q_2 raises the output resistance of the current source by the factor $g_{m2}r_{o2}$, which is the intrinsic voltage gain of Q_2. For GaAs MESFETs, $g_{m2}r_{o2}$ is typically 10 to 40. To allow a wide range of voltages at the output of the cascode current source, V_{BIAS} should be the lowest value that results in Q_1 operating in saturation.

EXERCISE

D6.24 For the cascode current source of Fig. 6.40 let $V_{SS} = 5$ V, $W_1 = 10$ μm, and $W_2 = 20$ μm, and assume that the devices have the typical parameter values given in Table 5.2. (a) Find the value of V_{BIAS} that will result in Q_1 operating at the edge of the saturation region (i.e., $V_{DS1} = |V_t|$) when the voltage at the output is –3 V. (b) What is the lowest allowable voltage at the current-source output? (c) What value of output current is obtained for $V_O = -3$ V? (d) What is the output resistance of the current source? (e) What change in output current results when the output voltage is raised from –3 V to +1 V?

Ans. (a) –4.3 V; (b) –3.3 V; (c) 1.1 mA; (d) 310 kΩ; (e) 0.013 mA

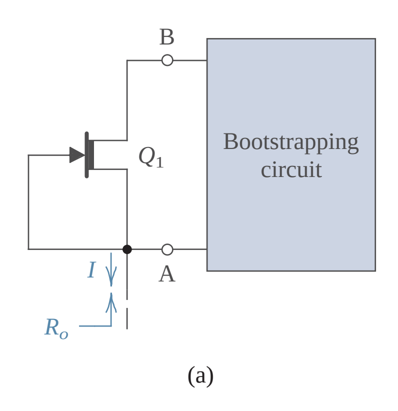
Increasing the Output Resistance by Bootstrapping

Another technique frequently employed to increase the effective output resistance of a MESFET, including the current-source-connected MESFET, is known as **bootstrapping.** The bootstrapping idea is illustrated in Fig. 6.41(a). Here the circuit inside the box senses the voltage at the bottom node of the current source, v_A, and causes a voltage v_B to appear at the top node of a value

$$v_B = V_S + \alpha v_A \tag{6.130}$$

where V_S is the dc voltage required to operate the current-source transistor in saturation $(V_S \geq |V_t|)$ and α is a constant ≤1. The incremental output resistance of the bootstrapped current source can be found by causing the voltage v_A to increase by an increment v_a. From Eq. (6.130) we find that the resulting increment in v_B is $v_b = \alpha v_a$. The incremental current through the current source is therefore $(v_a - v_b)/r_o$ or $(1-\alpha)v_a/r_o$. Thus the output resistance R_o is

$$R_o = \frac{v_a}{(1-\alpha)v_a/r_o} = \frac{r_o}{1-\alpha} \tag{6.131}$$



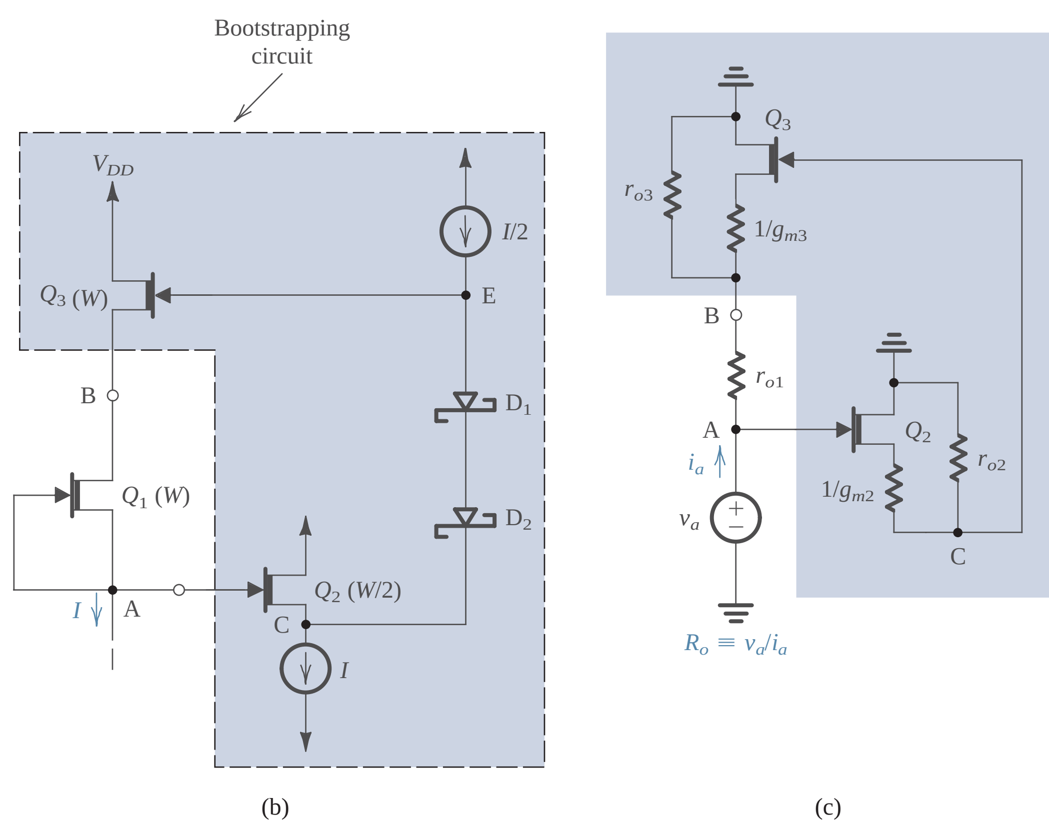


FIGURE 6.41 Bootstrapping of a MESFET current source Q_1: **(a)** basic arrangement; **(b)** an implementation; **(c)** small-signal equivalent circuit model of the circuit in (b), for the purpose of determining the output resistance R_o.

Thus, bootstrapping increases the output resistance by the factor $1/(1 - \alpha)$, which increases as α approaches unity. Perfect bootstrapping is achieved with $\alpha = 1$, resulting in $R_o = \infty$.

From the above we observe that the bootstrapping circuit senses whatever change occurs in the voltage at one terminal of the current source and causes an almost equal change to occur at the other terminal, thus maintaining an almost constant voltage across the current source and minimizing the change in current through the current-source transistor. The

action of the bootstrapping circuit can be likened to that of a person who attempts to lift himself off the ground by pulling on the straps of his boots (!), the origin of the name of this circuit technique, which, incidentally, predates GaAs technology. Bootstrapping is a form of positive feedback; the signal v_b that is fed back by the bootstrapping circuit is in phase with (has the same polarity as) the signal that is being sensed, v_a. Feedback will be studied formally in Chapter 8.

An implementation of the bootstrapped current source is shown in Fig. 6.41(b). Here transistor Q_2 is a source follower used to buffer node A, whose voltage is being sensed. The width of Q_2 is half that of Q_1 and is operating at half the bias current. (Transistors Q_1 and Q_2 are said to operate at the same **current density.**) Thus V_{GS} of Q_2 will be equal to that of Q_1—namely, zero—and hence $V_C = V_A$. The two Schottky diodes behave as a battery of approximately 1.4 V, resulting in the dc voltage at node E being 1.4 V higher than V_C. Note that the signal voltage at node C appears intact at node E; only the dc level is shifted. The diodes are said to perform **level shifting,** a common application of Schottky diodes in GaAs MESFET technology.

Transistor Q_3 is a source follower that is operating at the same current density as Q_1, and thus its V_{GS} must be zero, resulting in $V_B = V_E$. The end result is that the bootstrapping circuit causes a dc voltage of 1.4 V to appear across the current-source transistor Q_1. Provided that $|V_t|$ of Q_1 is less than 1.4 V, Q_1 will be operating in saturation as required.

To determine the output resistance of the bootstrapped current source, apply an incremental voltage v_a to node A, as shown in Fig. 6.41(c). Note that this small-signal equivalent circuit is obtained by implicitly using the T model (including r_o) for each FET and assuming that the Schottky diodes act as a perfect level shifter (that is, as an ideal dc voltage of 1.4 V with zero internal resistance). Analysis of this circuit is straightforward and yields

$$\alpha \equiv \frac{v_b}{v_a} = \frac{g_{m3}r_{o3}\dfrac{g_{m2}r_{o2}}{g_{m2}r_{o2}+1} + \dfrac{r_{o3}}{r_{o1}}}{g_{m3}r_{o3} + \dfrac{r_{o3}}{r_{o1}} + 1} \tag{6.132}$$

which is smaller than, but close to, unity, as required. The output resistance R_o is then obtained as

$$\begin{aligned} R_o \equiv \frac{v_a}{i_a} &= \frac{r_{o1}}{1-\alpha} \\ &= r_{o1}\frac{g_{m3}r_{o3} + (r_{o3}/r_{o1}) + 1}{g_{m3}r_{o3}/(g_{m2}r_{o2}+1) + 1} \end{aligned} \tag{6.133}$$

For $r_{o3} = r_{o1}$, assuming that $g_{m3}r_{o3}$ and $g_{m2}r_{o2}$ are $\gg 1$, and using the relationships for g_m and r_o for Q_2 and Q_3, one can show that

$$R_o \simeq r_{o1}(g_{m3}r_{o3}/2) \tag{6.134}$$

which represents an increase of about an order of magnitude in output resistance. Unfortunately, however, the circuit is rather complex.

A Simple Cascode Configuration–The Composite Transistor

The rather low output resistance of the MESFET places a severe limitation on the performance of MESFET current sources and various MESFET amplifiers. This problem can be alleviated by using the composite MESFET configuration shown in Fig. 6.42(a) in place of a

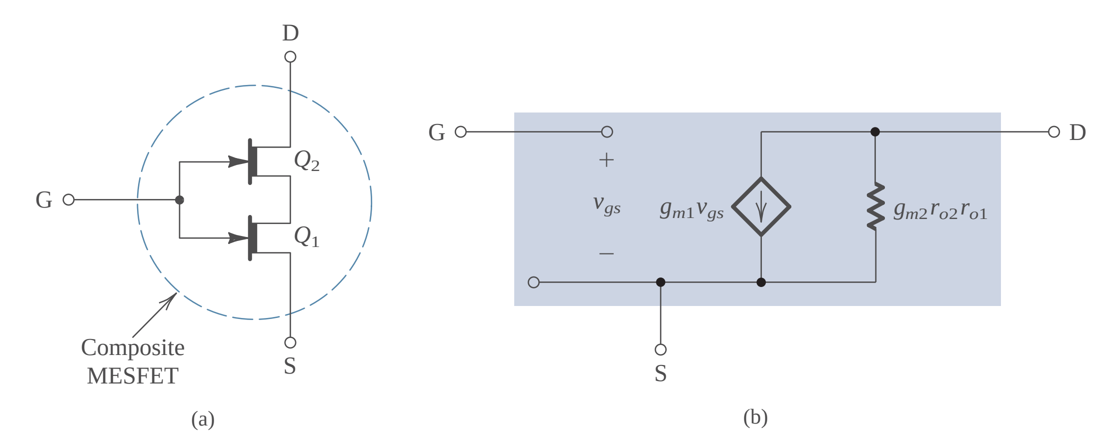


FIGURE 6.42 (**a**) The composite MESFET and (**b**) its small-signal model.

single MESFET. This circuit is unique to GaAs MESFETs and works only because of the early-saturation phenomenon observed in these devices. Recall from the discussion in Section 5.12 that **early saturation** refers to the fact that in a GaAs MESFET the drain current saturates at a voltage $v_{DS\text{sat}}$ that is lower than $v_{GS} - V_t$.

In the composite MESFET of Fig. 6.42(a), Q_2 is made much wider than Q_1. It follows that since the two devices are conducting the same current, Q_2 will have a gate-to-source voltage v_{GS2} whose magnitude is much closer to $|V_t|$ than $|v_{GS1}|$ is (thus, $|v_{GS2}| \gg |v_{GS1}|$). For instance, if we use the devices whose typical parameters are given in Table 5.2 and ignore for the moment channel-length modulation ($\lambda = 0$), we find that for $W_1 = 10$ μm and $W_2 = 90$ μm, at a current of 1 mA, $v_{GS1} = 0$ and $v_{GS2} = -\frac{2}{3}$ V. Now, since the drain-to-source voltage of Q_1 is $v_{DS1} = -v_{GS2} + v_{GS1}$, we see that v_{DS1} will be positive and close to but lower than $v_{GS1} - V_t$ ($\frac{2}{3}$ V in our example compared to 1 V). Thus in the absence of early saturation, Q_1 would be operating in the triode region. With early saturation, however, it has been found that saturation-mode operation is achieved for Q_1 by making Q_2 5 to 10 times wider.

The composite MESFET of Fig. 6.42(a) can be thought of as a cascode configuration, in which Q_2 is the cascode transistor, but without a separate bias line to feed the gate of the cascode transistor (as in Fig. 6.40). By replacing each of Q_1 and Q_2 with their small-signal models one can show that the composite device can be represented with the equivalent circuit model of Fig. 6.42(b). Thus while g_m of the composite device is equal to that of Q_1, the output resistance is increased by the intrinsic gain of Q_2, $g_{m2}r_{o2}$, which is typically in the range 10 to 40. This is a substantial increase and is the reason for the attractiveness of the composite MESFET.

The composite MESFET can be employed in any of the applications that can benefit from its increased output resistance. Some examples are shown in Fig. 6.43. The circuit in Fig. 6.43(a) is that of a current source with increased output resistance. Another view of the operation of this circuit can be obtained by considering Q_2 as a source follower that causes the drain of Q_1 to follow the voltage changes at the current-source terminal (node A), thereby bootstrapping Q_1 and increasing the effective output resistance of the current source. This alternative interpretation of circuit operation has resulted in its alternative name: the **self-bootstrapped** current source.

The application of the composite MESFET as a source follower is depicted in Fig. 6.43(b). Assuming the bias-current source I to be ideal, we can write for the gain of

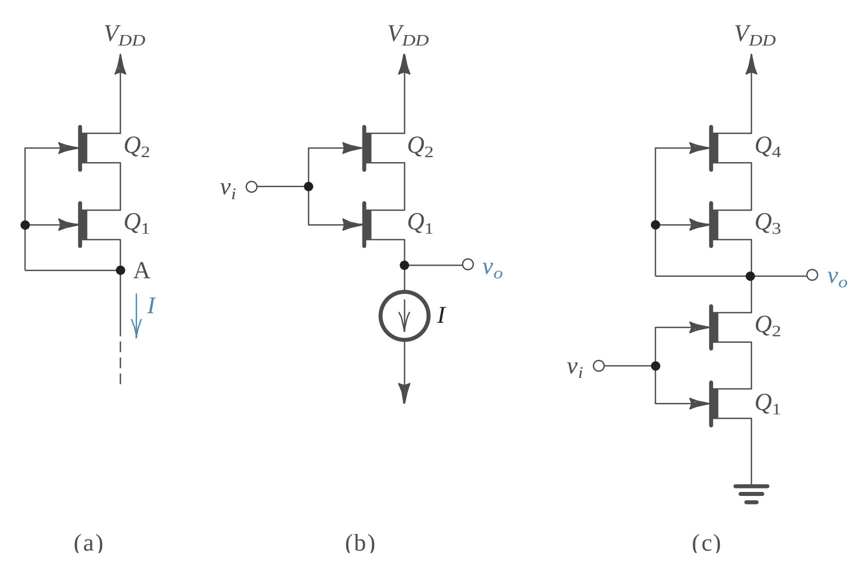


FIGURE 6.43 Applications of the composite MESFET: **(a)** as a current source; **(b)** as a source follower; and **(c)** as a gain stage.

this follower

$$\frac{v_o}{v_i} = \frac{r_{o,\text{eff}}}{r_{o,\text{eff}} + (1/g_{m1})}$$

$$= \frac{g_{m2} r_{o2} r_{o1}}{g_{m2} r_{o2} r_{o1} + (1/g_{m1})} \tag{6.135}$$

which is much closer to the ideal value of unity than is the gain of a single MESFET source follower.

EXERCISE

6.25 Using the device data given in Table 5.2, contrast the voltage gain of a source follower formed using a single MESFET having $W = 10\ \mu\text{m}$ with a composite MESFET follower with $W_1 = 10\ \mu\text{m}$ and $W_2 = 90\ \mu\text{m}$. In both cases assume biasing at 1 mA and neglect λ while calculating g_m (for simplicity).

Ans. Single: 0.952 V/V; composite: 0.999 V/V

A final example of the application of the composite MESFET is shown in Fig. 6.43(c). The circuit is a gain stage utilizing a composite MESFET (Q_1, Q_2) as a driver and another composite MESFET (Q_3, Q_4) as a current-source load. The small-signal gain is given by

$$\frac{v_o}{v_i} = -g_{m1} R_o \tag{6.136}$$

where R_o is the output resistance,

$$R_o = r_{o,\text{eff}}(Q_1, Q_2) // r_{o,\text{eff}}(Q_3, Q_4)$$

$$= g_{m2} r_{o2} r_{o1} // g_{m4} r_{o4} r_{o3} \tag{6.137}$$

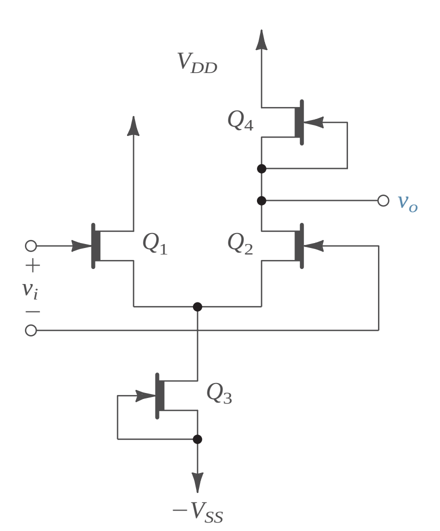


FIGURE 6.44 A simple MESFET differential amplifier.

Differential Amplifiers

The simplest possible implementation of a differential amplifier in GaAs MESFET technology is shown in Fig. 6.44. Here Q_1 and Q_2 form the differential pair, Q_3 forms the bias current source, and Q_4 forms the active (current-source) load. The performance of the circuit is impaired by the low output resistances of Q_3 and Q_4. The voltage gain is given by

$$\frac{v_o}{v_i} = -g_{m2}(r_{o2} // r_{o4}) \tag{6.138}$$

The gain can be increased by using one of the improved current-source implementations discussed above. Also, a rather ingenious technique has been developed for enhancing the gain of the MESFET differential pair. The circuit is shown in Fig. 6.45(a). While the drain

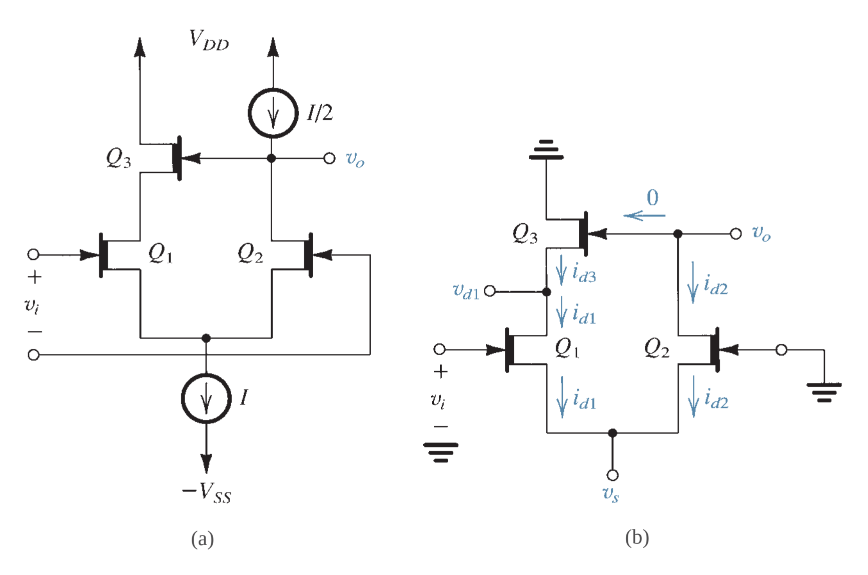


FIGURE 6.45 **(a)** A MESFET differential amplifier whose gain is enhanced by the application of positive feedback through the source follower Q_3; **(b)** small-signal analysis of the circuit in (a).

of Q_2 is loaded with a current-source load (as before), the output signal developed is fed back to the drain of Q_1 via the source follower Q_3. The small-signal analysis of the circuit is illustrated in Fig. 6.45(b) where the current sources I and $I/2$ have been assumed ideal and thus replaced with open circuits. To determine the voltage gain, we have grounded the gate terminal of Q_2 and applied the differential input signal v_i to the gate of Q_1. The analysis proceeds along the following steps:

1. From the output node we see that $i_{d2} = 0$.
2. From the sources node, since $i_{d2} = 0$, we find that $i_{d1} = 0$.
3. From the node at the drain of Q_1, since $i_{d1} = 0$, we find that $i_{d3} = 0$.
4. Writing for each transistor

$$i_d = g_m v_{gs} + v_{ds}/r_o = 0$$

we obtain three equations in the three unknowns v_{d1}, v_s, and v_o. The solution yields

$$\frac{v_o}{v_i} = g_{m1}r_{o1} \Big/ \left[\frac{g_{m1}r_{o1}+1}{g_{m2}r_{o2}+1} - \frac{g_{m3}r_{o3}}{g_{m3}r_{o3}+1}\right] \tag{6.139}$$

If all three transistors have the same geometry and are operating at equal dc currents, their g_m and r_o values will be equal and the expression in Eq. (6.139) reduces to

$$\frac{v_o}{v_i} \simeq (g_m r_o)^2 \tag{6.140}$$

Thus application of positive feedback through follower Q_3 enables one to obtain a gain equal to the square of that naturally available from a single stage!

EXERCISE

6.26 Using the device data given in Table 5.2, find the gain of the differential amplifier circuit of Fig. 6.45(a) for $I = 10$ mA and $W_1 = W_2 = W_3 = 100\ \mu$m.

Ans. 784 V/V



14.8 GALLIUM-ARSENIDE DIGITAL CIRCUITS

We conclude our study of digital-circuit families with a discussion of logic circuits implemented using the emerging technology of gallium arsenide. An introduction to this technology and its two basic devices, the MESFET and the Schottky-barrier diode (SBD), was given in Section 5.12. We urge the reader to review Section 5.12 before proceeding with the study of this section.

The major advantage that GaAs technology offers is a higher speed of operation than currently achievable using silicon devices. Gate delays of 10 to 100 ps have been reported for GaAs circuits. The disadvantages are a relatively high power dissipation per gate (1 to 10 mW); relatively small voltage swings and, correspondingly, narrow noise margins; low packing density, mostly as a result of the high-power dissipation per gate; and low manufacturing yield. The present state of affairs is that a few specialized manufacturers produce SSI, MSI, and some LSI digital circuits performing relatively specialized functions, with a cost per gate considerably higher than that of silicon digital ICs. Nevertheless, the very high

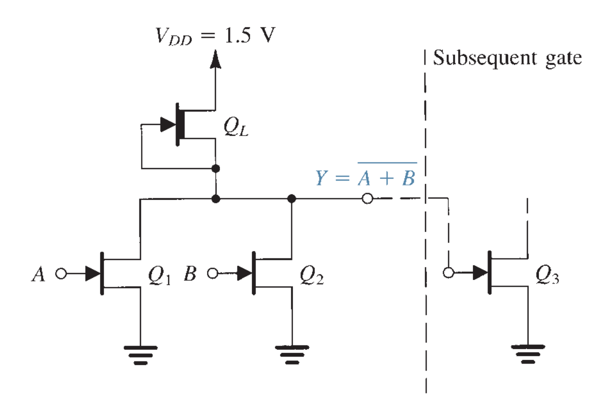


FIGURE 14.47 A DCFL GaAs gate implementing a two-input NOR function. The gate is shown driving the input transistor Q_3 of another gate.

speeds of operation achievable in GaAs circuits make it a worthwhile technology whose applications will possibly grow.

Unlike the CMOS logic circuits that we have studied in Chapter 13, and the bipolar logic families that we have studied in earlier sections of this chapter, there are no standard GaAs logic-circuit families. The lack of standards extends not only to the topology of the basic gates but also to the power-supply voltages used. In the following we present examples of the most popular GaAs logic gate circuits.

Direct-Coupled FET Logic (DCFL)

Direct-coupled FET logic (DCFL) is the simplest form of GaAs digital logic circuits. The basic gate is shown in Fig. 14.47. The gate utilizes enhancement MESFETs, Q_1 and Q_2, for the input switching transistors, and a depletion MESFET for the load transistor Q_L. The gate closely resembles the now obsolete depletion-load MOSFET circuit. The GaAs circuit of Fig. 14.47 implements a two-input NOR function.

To see how the MESFET circuit of Fig. 14.47 operates, ignore input B and consider the basic inverter formed by Q_1 and Q_L. When the input voltage applied to node A, v_I, is lower than the threshold voltage of the enhancement MESFET Q_1, denoted V_{tE}, transistor Q_1 will be off. Recall that V_{tE} is positive and for GaAs MESFETs is typically 0.1 to 0.3 V. Now if the gate output Y is open circuited, the output voltage will be very close to V_{DD}. In practice, however, the gate will be driving another gate, as indicated in Fig. 14.47, where Q_3 is the input transistor of the subsequent gate. In such a case, current will flow from V_{DD} through Q_L and into the gate terminal of Q_3. Recalling that the gate to source of a GaAs MESFET is a Schottky-barrier diode that exhibits a voltage drop of about 0.7 V when conducting, we see that the gate conduction of Q_3 will clamp the output high voltage (V_{OH}) to about 0.7 V. This is in sharp contrast to the MOSFET case, where no gate conduction takes place.

Figure 14.48 shows the DCFL inverter under study with the input of the subsequent gate represented by a Schottky diode Q_3. With $v_I < V_{tE}$, $i_1 = 0$ and i_L flows through Q_3 resulting in $v_O = V_{OH} \simeq 0.7$ V. Since V_{DD} is usually low (1.2 to 1.5 V) and the threshold voltage of Q_L, V_{tD}, is typically –0.7 to –1 V, Q_L will be operating in the triode region. (To simplify matters, we shall ignore in this discussion the early-saturation effect exhibited by GaAs MESFETs.)

As v_I is increased above V_{tE}, Q_1 turns on and conducts a current denoted i_1. Initially, Q_1 will be in the saturation region. Current i_1 subtracts from i_L, thus reducing the current in Q_3. The voltage across Q_3, v_O, decreases slightly. However, for the present discussion we shall assume that v_O will remain close to 0.7 V as long as Q_3 is conducting. This will continue until v_I reaches the value that results in $i_1 = i_L$. At this point, Q_3 ceases conduction

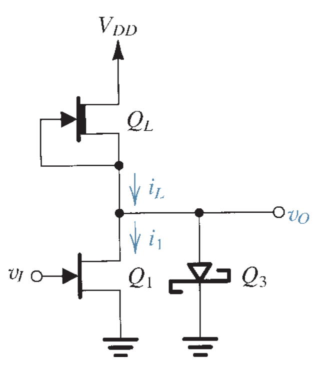


FIGURE 14.48 The DCFL gate with the input of the subsequent gate represented by a Schottky diode Q_3.

and can be ignored altogether. Further increase in v_I results in i_1 increasing, v_O decreasing, and $i_L = i_1$. When $(V_{DD} - v_O)$ exceeds $|V_{tD}|$, Q_L saturates; and when v_O falls below v_I by V_{tE}, Q_1 enters the triode region. Eventually, when $v_I = V_{OH} = 0.7$ V, $v_O = V_{OL}$, which is typically 0.1 to 0.2 V.

From the description above we see that the output voltage swing of the DCFL gate is limited by gate conduction to a value less than 0.7 V (typically 0.5 V or so). Further details on the operation of the DCFL gate are illustrated by the following example.

EXAMPLE 14.3

Consider a DCFL gate fabricated in a GaAs technology for which $L = 1$ μm, $V_{tD} = -1$ V, $V_{tE} =$ 0.2 V, β (for 1-μm width) $= 10^{-4}$ A/V^2, and $\lambda = 0.1$ V^{-1}. Let the widths of the input MESFETs be 50 μm, and let the width of the load MESFET be 6 μm. $V_{DD} = 1.5$ V. Using a constant-voltage-drop model for the gate-source Schottky diode with $V_D = 0.7$ V, and neglecting the early-saturation effect of GaAs MESFETs (that is, using Eqs. 5.120 to describe MESFET operation), find V_{OH}, V_{OL}, V_{IH}, NM_H, NM_L, the static power dissipation, and the propagation delay for a total equivalent capacitance at the gate output of 30 fF.

Solution

From the description above of the operation of the DCFL gate we found that $V_{OH} = 0.7$ V. To obtain V_{OL}, we consider the inverter in the circuit of Fig. 14.48 and let $v_I = V_{OH} = 0.7$ V. Since we expect $v_O = V_{OL}$ to be small, we assume Q_1 to be in the triode region and Q_L to be in saturation. (Q_3 is of course off.) Equating i_1 and i_L gives the equation

$$\beta_1[2(0.7-0.2)V_{OL} - V_{OL}^2](1 + 0.1V_{OL}) = \beta_L[0-(-1)]^2[1 + 0.1(1.5 - V_{OL})]$$

To simplify matters, we neglect the terms $0.1V_{OL}$ and substitute $\beta_L/\beta_1 = W_L/W_1 = 6/50$ to obtain a quadratic equation in V_{OL} whose solution gives $V_{OL} \simeq 0.17$ V.

Toward obtaining the value of V_{IL} we shall first find the value of v_I at which $i_1 = i_L$, the diode Q_3 turns off, and v_O begins to decrease. Since at this point $v_O = 0.7$ V, we assume that Q_1 is in saturation. Transistor Q_L has a v_{DS} of 0.8 V, which is less than $|V_{tD}|$ and is thus in the triode region. Equating i_1 and i_L gives

$$\beta_1(v_I - 0.2)^2(1 + 0.1 \times 0.7) = \beta_L[2(1)(1.5-0.7) - (1.5-0.7)^2][1 + 0.1(1.5-0.7)]$$

Substituting $\beta_L/\beta_1 = W_L/W_1 = 6/50$ and solving the resulting equation yields $v_I = 0.54$ V. Figure 14.49 shows a sketch of the transfer characteristic of the inverter. The slope dv_O/dv_I at

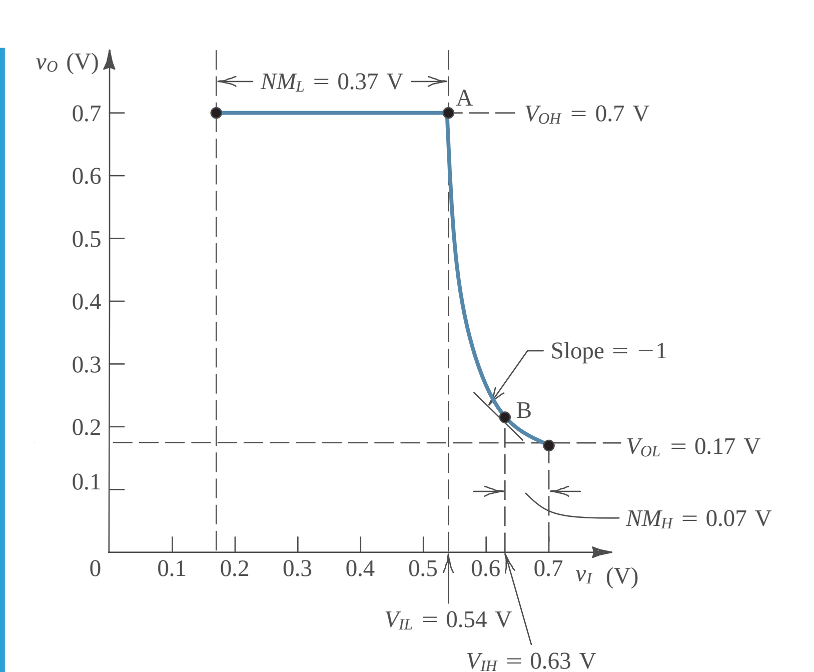


FIGURE 14.49 Transfer characteristic of the DCFL inverter of Fig. 14.48.

point A can be found to be –14.2 V/V. We shall consider point A as the point at which the inverter begins to switch from the high-output state; thus $V_{IL} \simeq 0.54$ V.

To obtain V_{IH}, we find the co-ordinates of point B at which $dv_O/dv_I = -1$. This can be done using a procedure similar to that employed for the MOSFET inverters and assuming Q_1 to be in the triode region and Q_L to be in saturation. Neglecting terms in $0.1v_O$, the result is $V_{IH} \simeq 0.63$ V. The noise margins can now be found as

$$NM_H \equiv V_{OH} - V_{IH} = 0.7 - 0.63 = 0.07 \text{ V}$$

$$NM_L \equiv V_{IL} - V_{OL} = 0.54 - 0.17 = 0.37 \text{ V}$$

The static power dissipation is determined by finding the supply current I_{DD} in the output-high and the output-low cases. When the output is high (at 0.7 V), Q_L is in the triode region and the supply current is

$$I_{DD} = \beta_L[2(0+1)(1.5-0.7)-(1.5-0.7)^2][1+0.1(1.5-0.7)]$$

Substituting $\beta_L = 10^{-4} \times W_L = 0.6 \text{ mA/V}^2$ results in

$$I_{DD} = 0.61 \text{ mA}$$

When the output is low (at 0.17 V), Q_L is in saturation and the supply current is

$$I_{DD} = \beta_L(0+1)^2[1+0.1(1.5-0.17)] = 0.68 \text{ mA}$$

Thus the average supply current is

$$I_{DD} = \tfrac{1}{2}(0.61+0.68) = 0.645 \text{ mA}$$

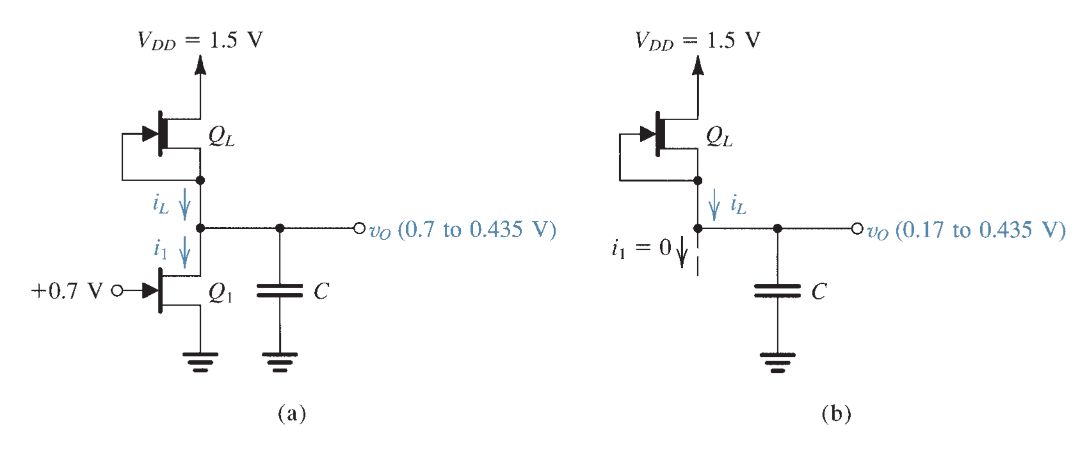


FIGURE 14.50 Circuits for calculating the propagation delays of the DCFL inverter: **(a)** t_{PHL}; **(b)** t_{PLH}.

and the static power dissipation is

$$P_D = 0.645 \times 1.5 \simeq 1 \text{ mW}$$

The propagation delay t_{PHL} is the time for the output voltage of the inverter to decrease from $V_{OH} = 0.7$ V to $\frac{1}{2}(V_{OH} + V_{OL}) = 0.435$ V. During this time v_I is at the high level of 0.7 V, and the capacitance C (assumed to be 30 fempto Farads [fF]) is discharged by $(i_1 - i_L)$; refer to Fig. 14.50(a). The average discharge current is found by calculating i_1 and i_L at the beginning and at the end of the discharge interval. The result is that i_1 changes from 1.34 mA to 1.28 mA and i_L changes from 0.61 mA to 0.66 mA. Thus the discharge current $(i_1 - i_L)$ changes from 0.73 mA to 0.62 mA for an average value of 0.675 mA. Thus

$$t_{PHL} = \frac{C\Delta V}{I} = \frac{30 \times 10^{-15}(0.7 - 0.435)}{0.675 \times 10^{-3}} = 11.8 \text{ ps}$$

To determine t_{PLH} we refer to the circuit in Fig. 14.50(b) and note that during t_{PLH}, v_O changes from $V_{OL} = 0.17$ V to $\frac{1}{2}(V_{OH} + V_{OL}) = 0.435$ V. The charging current is the average value of i_L, which changes from 0.8 mA to 0.66 mA. Thus $i_{L|\text{average}} = 0.73$ mA and

$$t_{PLH} = \frac{30 \times 10^{-15} \times (0.435 - 0.17)}{0.73 \times 10^{-3}} = 10.9 \text{ ps}$$

The propagation delay of the DCFL gate can now be found as

$$t_P = \tfrac{1}{2}(t_{PHL} + t_{PLH}) = 11.4 \text{ ps}$$

As a final remark, we note that the analysis above was done using simplified device models; our objective is to show how the circuit works rather than to find accurate performance measures. These can be obtained using SPICE simulation with more elaborate models [see Roberts and Sedra (1997)].

Logic Gates Using Depletion MESFETs

The DCFL circuits studied above require both enhancement and depletion devices and thus are somewhat difficult to fabricate. Also, owing to the fact that the voltage swings and noise margins are rather small, very careful control of the value of V_{tE} is required in fabrication. As an alternative, we now present circuits that utilize depletion devices only.

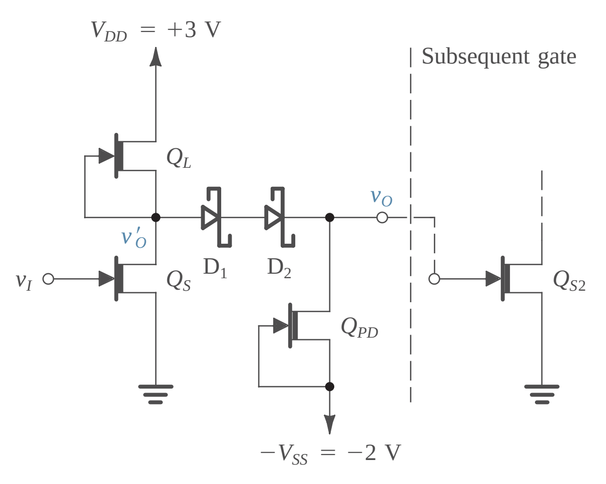


FIGURE 14.51 An inverter circuit utilizing depletion-mode devices only. Schottky diodes are employed to shift the output logic levels to values compatible with the input levels required to turn the depletion MESFET Q_S on and off. This circuit is known as FET logic (FL).

Figure 14.51 shows the basic inverter circuit of a family of GaAs logic circuits known at FET logic (FL). The heart of the FL inverter is formed by the switching transistor Q_S and its load Q_L—both depletion-type MESFETs. Since the threshold voltage of a depletion MESFET, V_{tD}, is negative, a negative voltage $<V_{tD}$ is needed to turn Q_S off. On the other hand the output low voltage at the drain of Q_S will always be positive. It follows that the logic levels at the drain of Q_S are not compatible with the levels required at the gate input. The incompatibility problem is solved by simply shifting the level of the voltage v'_O down by two diode drops, that is, by approximately 1.4 V. This level shifting is accomplished by the two Schottky diodes D_1 and D_2. The depletion transistor Q_{PD} provides a constant-current bias for D_1 and D_2. To ensure that Q_{PD} operates in the saturation region at all times, its source is connected to a negative supply $-V_{SS}$, and the value of V_{SS} is selected to be equal to or greater than the lowest level of $v_O(V_{OL})$ plus the magnitude of the threshold voltage, $|V_{tD}|$. Transistor Q_{PD} also supplies the current required to discharge a load capacitance when the output voltage of the gate goes low, hence the name "pull-down" transistor and the subscript *PD*.

To see how the inverter of Fig. 14.51 operates, refer to its transfer characteristic, shown in Fig. 14.52. The circuit is usually designed using MESFETs having equal channel lengths (typically 1 μm) and having widths $W_S = W_L = 2W_{PD}$. The transfer characteristic shown is for the case $V_{tD} = -0.9$ V. For v_I lower than V_{tD}, Q_S will be off and Q_L will operate in saturation, supplying a constant current I_L to D_1 and D_2. Transistor Q_{PD} will also operate in saturation with a constant current $I_{PD} = \frac{1}{2}I_L$. The difference between the two currents will flow through the gate terminal of the input transistor of the next gate in the chain, Q_{S2}. Thus the input Schottky diode of Q_{S2} clamps the output voltage v_O to approximately 0.7 V, which is the output high level, V_{OH}. (Note that for this discussion we shall neglect the finite output resistance in saturation.)

As v_I is raised above V_{tD}, Q_S turns on. Since its drain is at +2.1 V, Q_S will operate in the saturation region and will take away some of the current supplied by Q_L. Thus the current flowing into the gate of Q_{S2} decreases by an equal amount. If we keep increasing v_I, a value is reached for which the current in Q_S equals $\frac{1}{2}I_L$, thus leaving no current to flow through the gate of Q_{S2}. This corresponds to the point labeled A on the transfer characteristic. A further slight increase in v_I will cause the voltage v'_O to fall to the point B where Q_S enters the triode region. The segment AB of the transfer curve represents the high-gain region of operation, having a slope equal to $-g_{ms}R$ where R denotes the total equivalent resistance at the drain node. Note that this segment is shown as vertical in Fig. 14.52 because we are neglecting the output resistance in saturation.

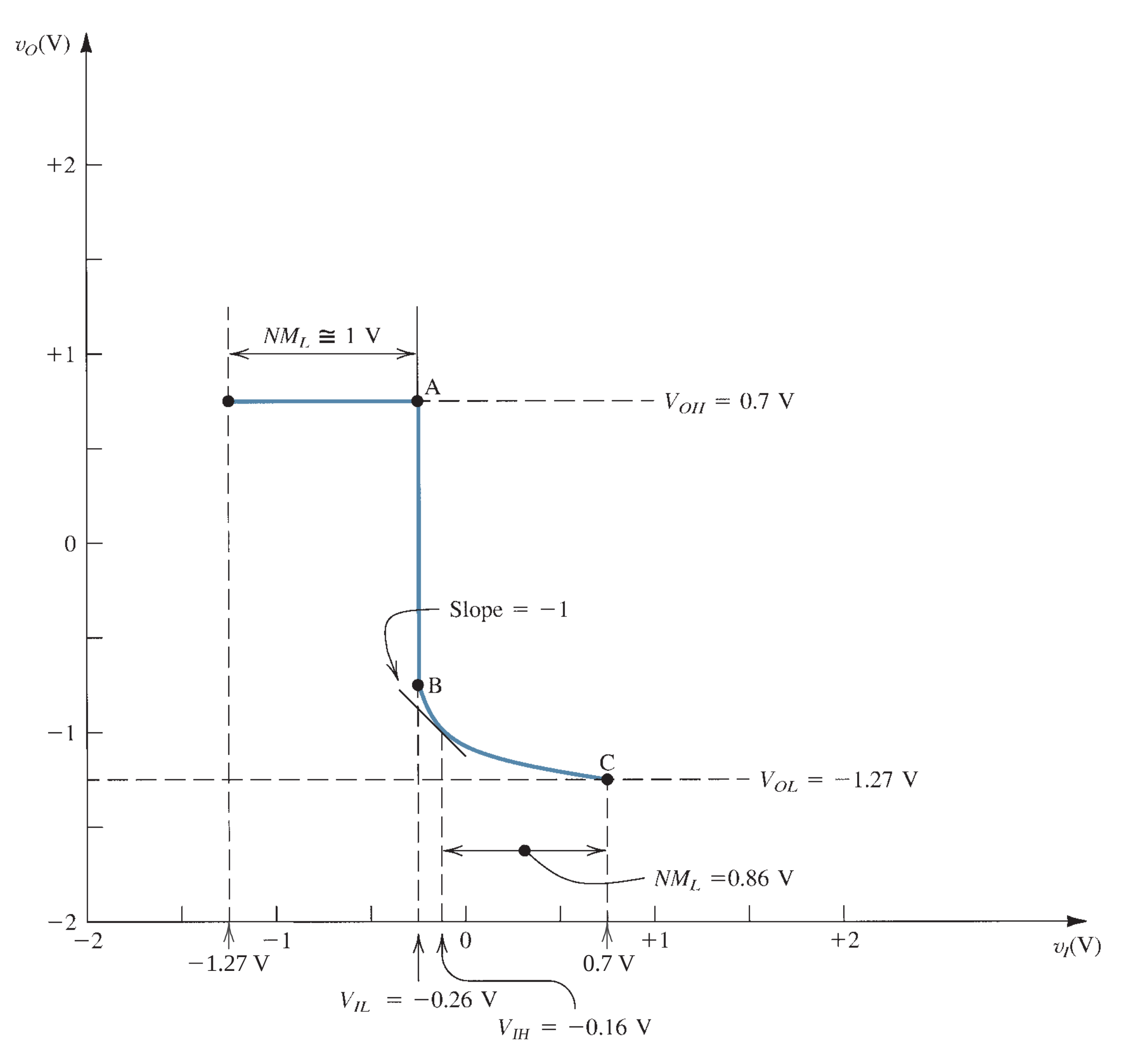


FIGURE 14.52 Transfer characteristic of the FL inverter of Fig. 14.51.

The segment BC of the transfer curve corresponds to Q_S operating in the triode region. Here Q_L and Q_{PD} continue to operate in saturation and D_1 and D_2 remain conducting. Finally, for $v_I = V_{OH} = 0.7$ V, $v_O = V_{OL}$, which for the case $V_{tD} = -0.9$ V can be found to be –1.3 V.

EXERCISE

14.30 Verify that the co-ordinates of points A, B, and C of the transfer characteristic are as indicated in Fig. 14.52. Let $V_{tD} = -0.9$ V and $\lambda = 0$.

As indicated in Fig. 14.52, the FL inverter exhibits much higher noise margins than those for the DCFL circuit. The FL inverter, however, requires two power supplies.

The FL inverter can be used to construct a NOR gate by simply adding transistors with drain and source connected in parallel with those of Q_S.

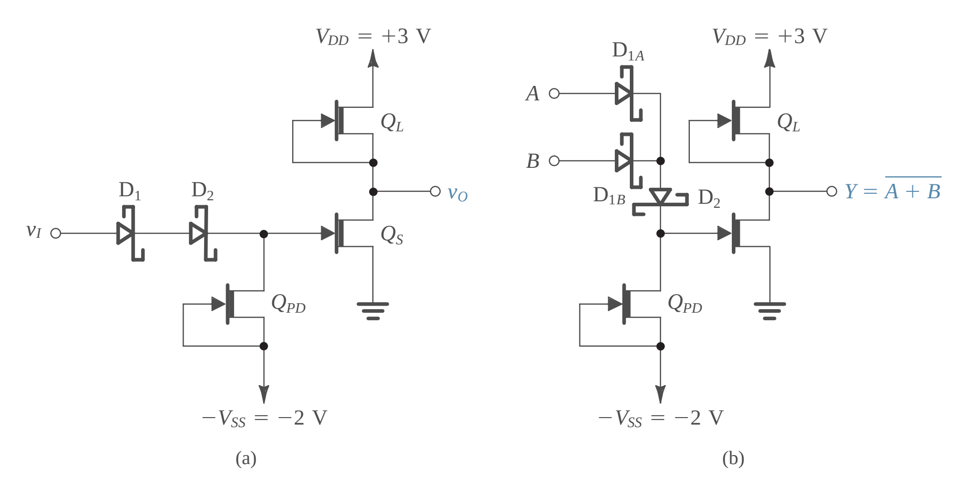


FIGURE 14.53 **(a)** An SDFL inverter. **(b)** An SDFL NOR gate.

Schottky Diode FET Logic (SDFL)

If the diode level-shifting network of the FL inverter is connected at the input side of the gate, rather than at the output side, we obtain the circuit shown in Fig. 14.53(a). This inverter operates in much the same manner as the FL inverter. The modified circuit, however, has a very interesting feature: The NOR function can be implemented by simply connecting additional diodes, as shown in Fig. 14.53(b). This logic form is known as Schottky diode FET logic (SDFL). SDFL permits higher packing density than other forms of MESFET logic because only an additional diode, rather than an additional transistor, is required for each additional input, and diodes require much smaller areas than transistors.

Buffered FET Logic (BFL)

Another variation on the basic FL inverter of Fig. 14.51 is possible. A source follower can be inserted between the drain of Q_S and the diode level-shifting network. The resulting gate, shown for the case of a two-input NOR, is depicted in Fig. 14.54. This form of GaAs logic circuit is known as buffered FET logic (BFL). The source-follower transistor Q_{SF} increases the output current-driving capability, thus decreasing the low-to-high propagation time. FL, BFL, and SDFL feature propagation delays of the order of 100 ps and power dissipation of the order of 10 mW/gate.

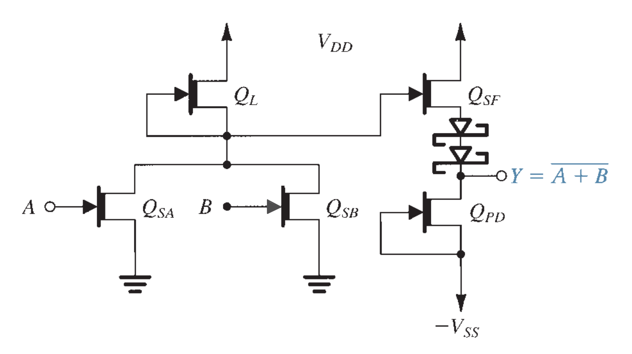


FIGURE 14.54 A BFL two-input NOR gate. The gate is formed by inserting a source-follower transistor Q_{SF} between the inverting stage and the level-shifting stage.



14.3 TRANSISTOR–TRANSISTOR LOGIC (TTL OR T^2L)

For more than two decades (late 1960s to late 1980s) TTL enjoyed immense popularity. Indeed, the bulk of digital systems applications employing SSI and MSI packages were designed using TTL.

We shall begin this section with a study of the evolution of TTL from DTL. In this way we shall explain the function of each of the stages of the complete TTL gate circuit. Characteristics of standard TTL gates will be studied in Section 14.4. Standard TTL, however, has now been virtually replaced with more advanced forms of TTL that feature improved performance. These will be discussed in Section 14.5.

Evolution of TTL from DTL

The basic DTL gate circuit in discrete form was discussed in the previous section (see Fig. 14.6). The integrated-circuit form of the DTL gate is shown in Fig. 14.7 with only one input indicated. As a prelude to introducing TTL, we have drawn the input diode as a diode-connected transistor (Q_1), which corresponds to how diodes are made in IC form.

This circuit differs from the discrete DTL circuit of Fig. 14.6 in two important aspects. First, one of the steering diodes is replaced by the base–emitter junction of a transistor (Q_2) that is either cut off (when the input is low) or in the active mode (when the input is high). This is done to reduce the input current and thereby increase the fan-out capability of the gate. A detailed explanation of this point, however, is not relevant to our study of TTL. Second, the resistance R_B is returned to ground rather than to a negative supply, as was done in the earlier discrete circuit. An obvious advantage of this is the elimination of the additional power supply. The disadvantage, however, is that the reverse base current available to remove the excess charge stored in the base of Q_3 is rather small. We shall elaborate on this point below.

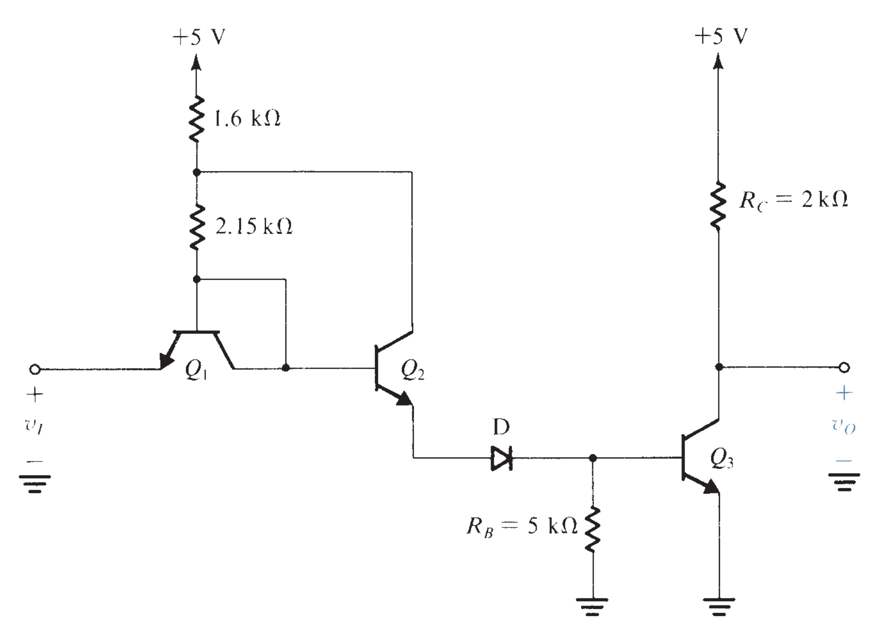


FIGURE 14.7 IC form of the DTL gate with the input diode shown as a diode-connected transistor (Q_1). Only one input terminal is shown.

EXERCISE

14.4 Consider the DTL gate circuit shown in Fig. 14.7 and assume that $\beta(Q_2) = \beta(Q_3) = 50$. (a) When $v_I =$ 0.2 V, find the input current. (b) When $v_I = +5$ V, find the base current of Q_3.

Ans. (a) 1.1 mA; (b) 1.6 mA

Reasons for the Slow Response of DTL

The DTL gate has relatively good noise margins and reasonably good fan-out capability. Its response, however, is rather slow. There are two reasons for this: first, when the input goes low and Q_2 and D turn off, the charge stored in the base of Q_3 has to leak through R_B to ground. The initial value of the reverse base current that accomplishes this "base discharging" process is approximately 0.7 V/R_B, which is about 0.14 mA. Because this current is quite small in comparison to the forward base current, the time required for the removal of base charge is rather long, which contributes to lengthening the gate delay.

The second reason for the relatively slow response of DTL derives from the nature of the output circuit of the gate, which is simply a common-emitter transistor. Figure 14.8 shows the output transistor of a DTL gate driving a capacitive load C_L. The capacitance C_L represents the input capacitance of another gate and/or the wiring and parasitic capacitances that are inevitably present in any circuit. When Q_3 is turned on, its collector voltage cannot instantaneously fall because of the existence of C_L. Thus Q_3 will not immediately saturate but rather will operate in the active region. The collector of Q_3 will therefore act as a constant-current source and will sink a relatively large current (βI_B). This large current will rapidly discharge C_L. We thus see that the common-emitter output stage features a short turn-on time. However, turnoff is another matter.

Consider next the operation of the common-emitter output stage when Q_3 is turned off. The output voltage will not rise immediately to the high level (V_{CC}). Rather, C_L will charge up to V_{CC} through R_C. This is a rather slow process, and it results in lengthening the DTL gate delay (and similarly the RTL gate delay).

Having identified the two reasons for the slow response of DTL, we shall see in the following how these problems are remedied in TTL.

Input Circuit of the TTL Gate

Figure 14.9 shows a conceptual TTL gate with only one input terminal indicated. The most important feature to note is that the input diode has been replaced by a transistor. One can

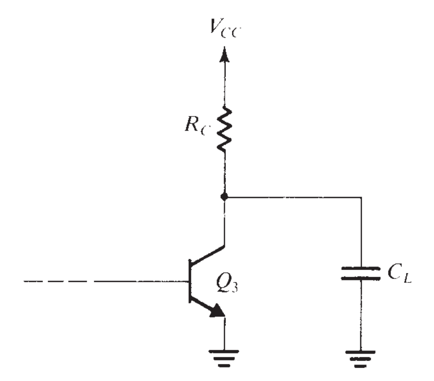


FIGURE 14.8 The output circuit of a DTL gate driving a capacitive load C_L.

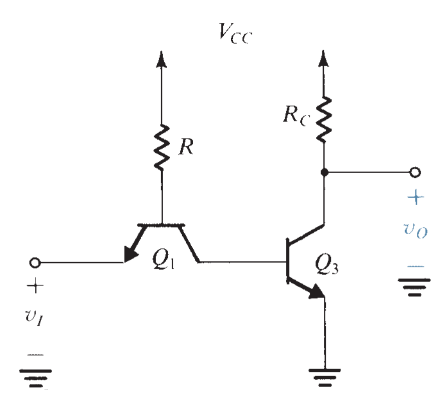


FIGURE 14.9 Conceptual form of TTL gate. Only one input terminal is shown.

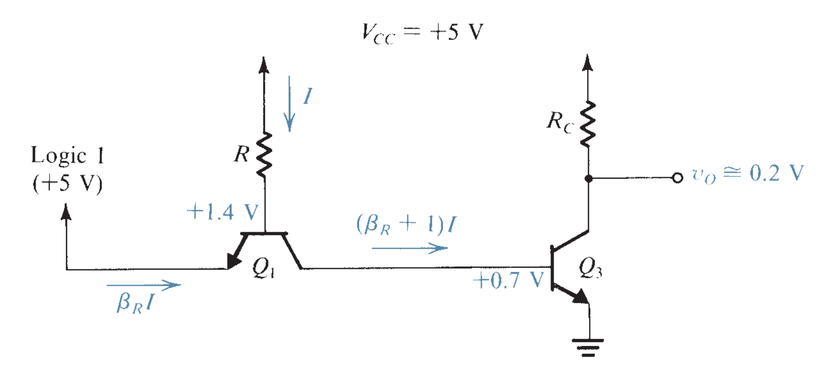


FIGURE 14.10 Analysis of the conceptual TTL gate when the input is high.

think of this simply as if the short circuit between base and collector of Q_1 in Fig. 14.7 has been removed.

To see how the conceptual TTL circuit of Fig. 14.9 works, let the input v_I be high (say, $v_I = V_{CC}$). In this case current will flow from V_{CC} through R, thus forward-biasing the base–collector junction of Q_1. Meanwhile, the base–emitter junction of Q_1 will be reverse-biased. Therefore Q_1 will be operating in the **inverse active mode**—that is, in the active mode but with the roles of emitter and collector interchanged. The voltages and currents will be as indicated in Fig. 14.10, where the current I can be calculated from

$$I = \frac{V_{CC} - 1.4}{R}$$

In actual TTL circuits Q_1 is designed to have a very low reverse β ($\beta_R \simeq 0.02$). Thus the gate input current will be very small, and the base current of Q_3 will be approximately equal to I. This current will be sufficient to drive Q_3 into saturation, and the output voltage will be low (0.1 to 0.2 V).

Next let the gate input voltage be brought down to the logic-0 level (say, $v_I \simeq 0.2$ V). The current I will then be diverted to the emitter of Q_1. The base–emitter junction of Q_1 will become forward-biased, and the base voltage of Q_1 will therefore drop to 0.9 V. Since Q_3 *was* in saturation, its base voltage will remain at +0.7 V pending the removal of the excess charge stored in the base region. Figure 14.11 indicates the various voltage and current values immediately after the input is lowered. We see that Q_1 will be operating in the normal

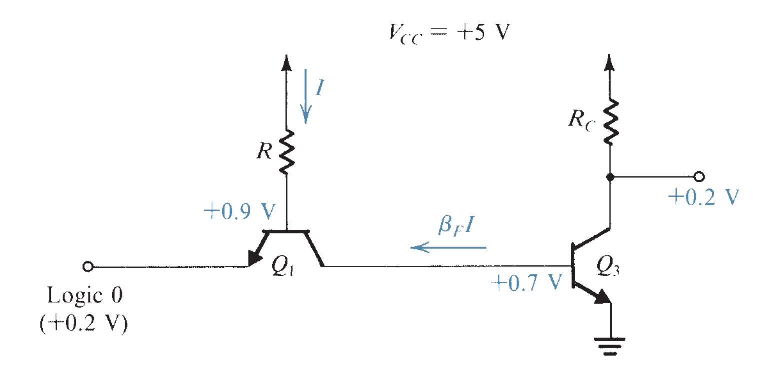


FIGURE 14.11 Voltage and current values in the conceptual TTL circuit immediately after the input voltage is lowered.

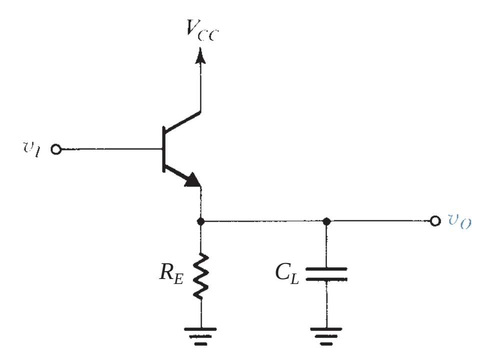


FIGURE 14.12 An emitter-follower output stage with capacitive load.

active mode[3] and its collector will carry a large current ($\beta_F I$). This large current rapidly discharges the base of Q_3 and drives it into cutoff. We thus see the action of Q_1 in speeding up the turn-off process.

As Q_3 turns off, the voltage at its base is reduced, and Q_1 enters the saturation mode. Eventually the collector current of Q_1 will become negligibly small, which implies that its V_{CEsat} will be approximately 0.1 V and the base of Q_3 will be at about 0.3 V, which keeps Q_3 in cutoff.

Output Circuit of the TTL Gate

The above discussion illustrates how one of the two problems that slow down the operation of DTL is solved in TTL. The second problem, the long rise time of the output waveform, is solved by modifying the output stage, as we shall now explain.

First, recall that the common-emitter output stage provides fast discharging of load capacitance but rather slow charging. The opposite is obtained in the emitter-follower output stage shown in Fig. 14.12. Here, as v_I goes high, the transistor turns on and provides a low output resistance (characteristic of emitter followers), which results in fast charging of C_L. On the other hand, when v_I goes low, the transistor turns off and C_L is then left to discharge slowly through R_E.

[3] Although the collector voltage of Q_1 is lower than its base voltage by 0.2 V, the collector–base junction will in effect be cut off and Q_1 will be operating in the active mode.

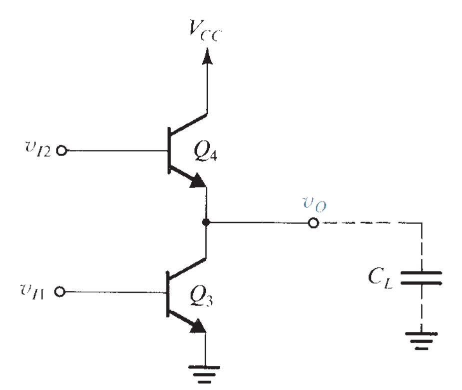


FIGURE 14.13 The totem-pole output stage.

It follows that an optimum output stage would be a combination of the common-emitter and the emitter-follower configurations. Such an output stage, shown in Fig. 14.13, has to be driven by two *complementary* signals v_{I1} and v_{I2}. When v_{I1} is high v_{I2} will be low, and in this case Q_3 will be on and saturated, and Q_4 will be off. The common-emitter transistor Q_3 will then provide the fast discharging of load capacitance and in steady state provide a low resistance (R_{CEsat}) to ground. Thus when the output is low, the gate can *sink* substantial amounts of current through the saturated transistor Q_3.

When v_{I1} is low and v_{I2} is high, Q_3 will be off and Q_4 will be conducting. The emitter follower Q_4 will then provide fast charging of load capacitance. It also provides the gate with a low output resistance in the high state and hence with the ability to *source* a substantial amount of load current.

Because of the appearance of the circuit in Fig. 14.13, with Q_4 stacked on top of Q_3, the circuit has been given the name **totem-pole output stage.** Also, because of the action of Q_4 in *pulling up* the output voltage to the high level, Q_4 is referred to as the **pull-up transistor.** Since the pulling up is achieved here by an active element (Q_4), the circuit is said to have an **active pull-up.** This is in contrast to the **passive pull-up** of RTL and DTL gates. Of course, the common-emitter transistor Q_3 provides the circuit with **active pull-down.** Finally, note that a special **driver circuit** is needed to generate the two complementary signals v_{I1} and v_{I2}.

EXAMPLE 14.1

We wish to analyze the circuit shown together with its driving waveforms in Fig. 14.14 to determine the waveform of the output signal v_O. Assume that Q_3 and Q_4 have $\beta = 50$.

Solution

Consider first the situation before v_{I1} goes high—that is, at time $t < 0$. In this case Q_3 is off and Q_4 is on, and the circuit can be simplified to that shown in Fig. 14.15. In this simplified circuit we have replaced the voltage divider (R_1, R_2) by its Thévenin equivalent. In the steady state, C_L will be charged to the output voltage v_O, whose value can be obtained as follows:

$$5 = 10 \times I_B + V_{BE} + I_E \times 0.5 + 2.5$$

Substituting $V_{BE} \simeq 0.7$ V and $I_B = I_E/(\beta + 1) = I_E/51$ gives $I_E = 2.59$ mA. Thus the output voltage v_O is given by

$$v_O = 2.5 + I_E \times 0.5 = 3.79 \text{ V}$$

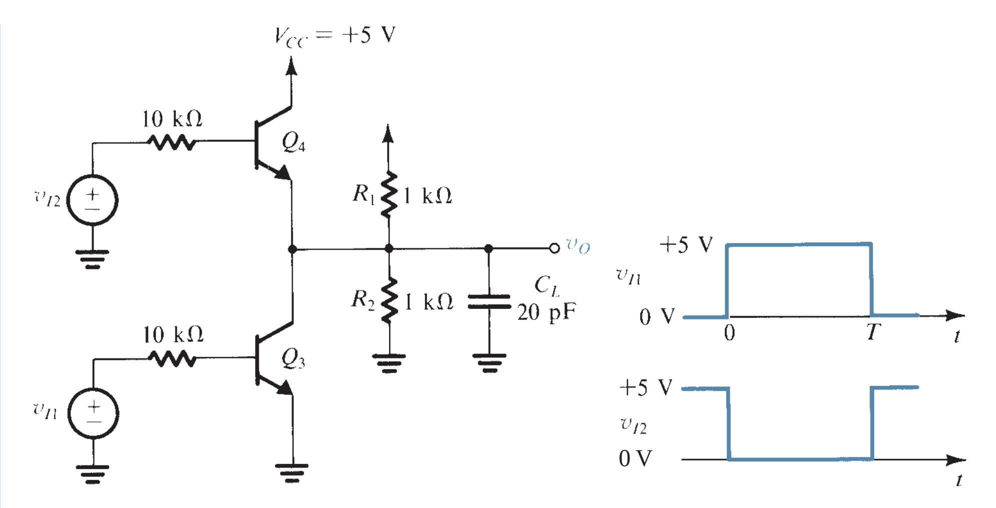


FIGURE 14.14 Circuit and input waveforms for Example 14.1.

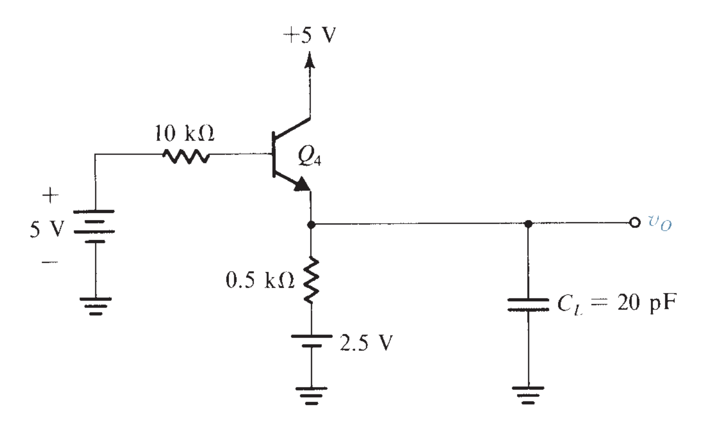


FIGURE 14.15 The circuit of Fig. 14.14 when Q_3 is off.

We next consider the circuit as v_{I1} goes high and v_{I2} goes low. Transistor Q_3 turns on and transistor Q_4 turns off, and the circuit simplifies to that shown in Fig. 14.16. Again we have used the Thévenin equivalent of the divider (R_1, R_2). We shall also assume that the switching times of the transistors are negligibly small. Thus at $t = 0+$ the base current of Q_3 becomes

$$I_B = \frac{5 - 0.7}{10} = 0.43 \text{ mA}$$

Since at $t = 0$ the collector voltage of Q_3 is 3.79 V, and since this value cannot change instantaneously because of C_L, we see that at $t = 0+$ transistor Q_3 will be in the active mode. The collector current of Q_3 will be βI_B, which is 21.5 mA, and the circuit will have the equivalent shown in Fig. 14.17(a). A simpler version of this equivalent circuit, obtained using Thévenin's theorem, is shown in Fig. 14.17(b).

The equivalent circuit of Fig. 14.17 applies as long as Q_3 remains in the active mode. This condition persists while C_L is being discharged and until v_O reaches about +0.3 V, at which time Q_3 enters saturation. This is illustrated by the waveform in Fig. 14.18. The time for the output

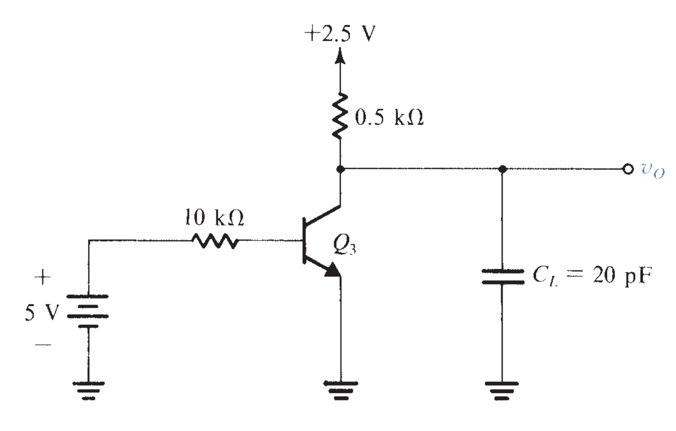


FIGURE 14.16 The circuit of Fig. 14.14 when Q_4 is off.

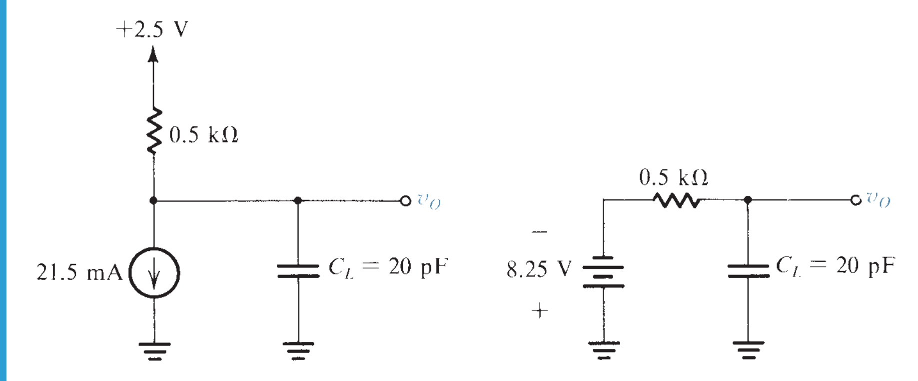


FIGURE 14.17 **(a)** Equivalent circuit for the circuit in Fig. 14.16 when Q_3 is in the active mode. **(b)** Simpler version of the circuit in (a) obtained using Thévenin's theorem.

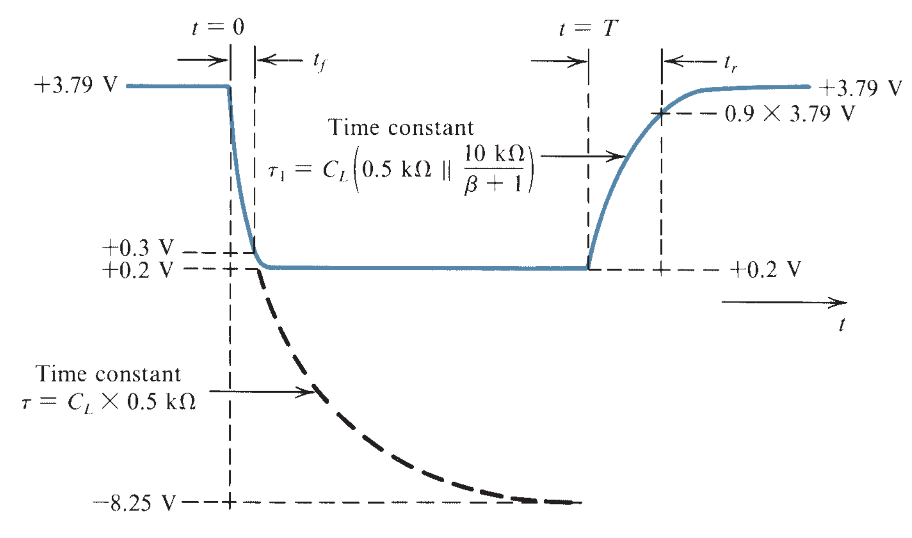


FIGURE 14.18 Details of the output voltage waveform for the circuit in Fig. 14.14.

voltage to fall from +3.79 V to +0.3 V, which can be considered the **fall time** t_f, can be obtained from

$$-8.25 - (-8.25 - 3.79)e^{-t_f/\tau} = 0.3$$

which results in

$$t_f \simeq 0.34\tau$$

where

$$\tau = C_L \times 0.5\text{ k}\Omega = 10\text{ ns}$$

Thus $t_f = 3.4$ ns.

After Q_3 enters saturation, the capacitor discharges further to the final steady-state value of V_{CEsat} ($\simeq$0.2 V). The transistor model that applies during this interval is more complex; since the interval in question is quite short, we shall not pursue the matter further.

Consider next the situation as v_{I1} goes low and v_{I2} goes high at $t = T$. Transistor Q_3 turns off as Q_4 turns on. We shall assume that this occurs immediately, and thus at $t = T+$ the circuit simplifies to that in Fig. 14.15. We have already analyzed this circuit in the steady state and thus know that eventually v_O will reach +3.79 V. Thus v_O rises exponentially from +0.2 V toward +3.79 V with a time constant of $C_L\{0.5\text{ k}\Omega//[10\text{ k}\Omega/(\beta + 1)]\}$, where we have neglected the emitter resistance r_e. Denoting this time constant τ_1, we obtain $\tau_1 = 2.8$ ns. Defining the rise time t_r as the time for v_O to reach 90% of the final value, we obtain $3.79 - (3.79 - 0.2)e^{-t_r/\tau_1} = 0.9 \times 3.79$, which results in $t_r = 6.4$ ns. Figure 14.18 illustrates the details of the output voltage waveform.

The Complete Circuit of the TTL Gate

Figure 14.19 shows the complete TTL gate circuit. It consists of three stages: the input transistor Q_1, whose operation has already been explained, the driver stage Q_2, whose function is to generate the two complementary voltage signals required to drive the totem-pole circuit,

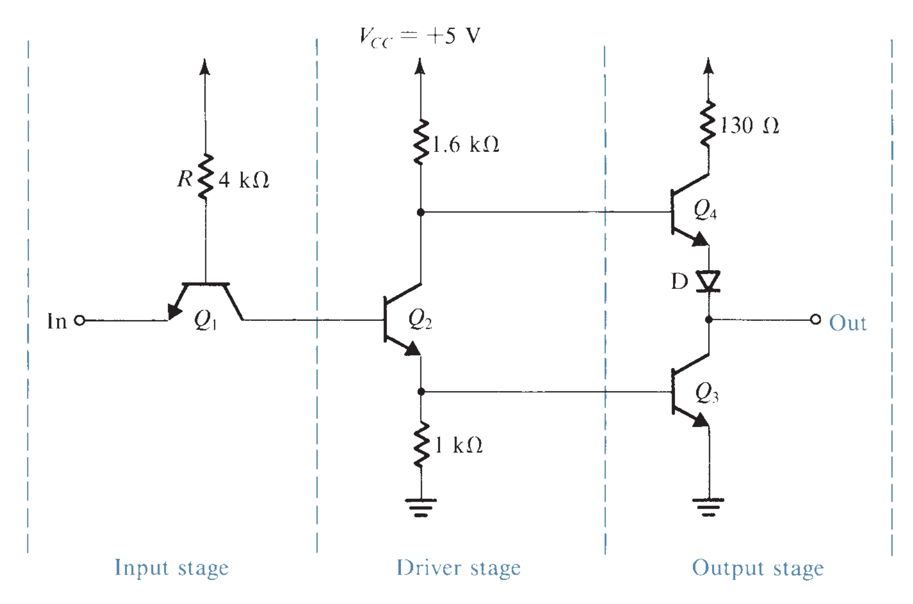


FIGURE 14.19 The complete TTL gate circuit with only one input terminal indicated.

which is the third (output) stage of the gate. The totem-pole circuit in the TTL gate has two additional components: the 130-Ω resistance in the collector circuit of Q_4 and the diode D in the emitter circuit of Q_4. The function of these two additional components will be explained shortly. Notice that the TTL gate is shown with only one input terminal indicated. Inclusion of additional input terminals will be considered in Section 14.4.

Because the driver stage Q_2 provides two complementary (that is, out-of-phase) signals, it is known as a **phase splitter.**

We shall now provide a detailed analysis of the TTL gate circuit in its two extreme states: one with the input high and one with the input low.

Analysis When the Input Is High

When the input is high (say, +5 V), the various voltages and currents of the TTL circuit will have the values indicated in Fig. 14.20. The analysis illustrated in Fig. 14.20 is quite straightforward, and the order of the steps followed is indicated by the circled numbers. As expected, the input transistor is operating in the inverse active mode, and the input current, called the **input high current** I_{IH}, is small; that is,

$$I_{IH} = \beta_R I \simeq 15\ \mu\text{A}$$

where we assume that $\beta_R \simeq 0.02$.

The collector current of Q_1 flows into the base of Q_2, and its value is sufficient to saturate the phase-splitter transistor Q_2. The latter supplies the base of Q_3 with sufficient current to drive it into saturation and lower its output voltage to $V_{CE\text{sat}}$ (0.1 to 0.2 V). The voltage at the collector of Q_2 is $V_{BE3} + V_{CE\text{sat}}(Q_2)$, which is approximately +0.9 V. If diode D were not included, this voltage would be sufficient to turn Q_4 on, which is contrary to the proper operation of the totem-pole circuit. Including diode D ensures that both Q_4 and D remain off.

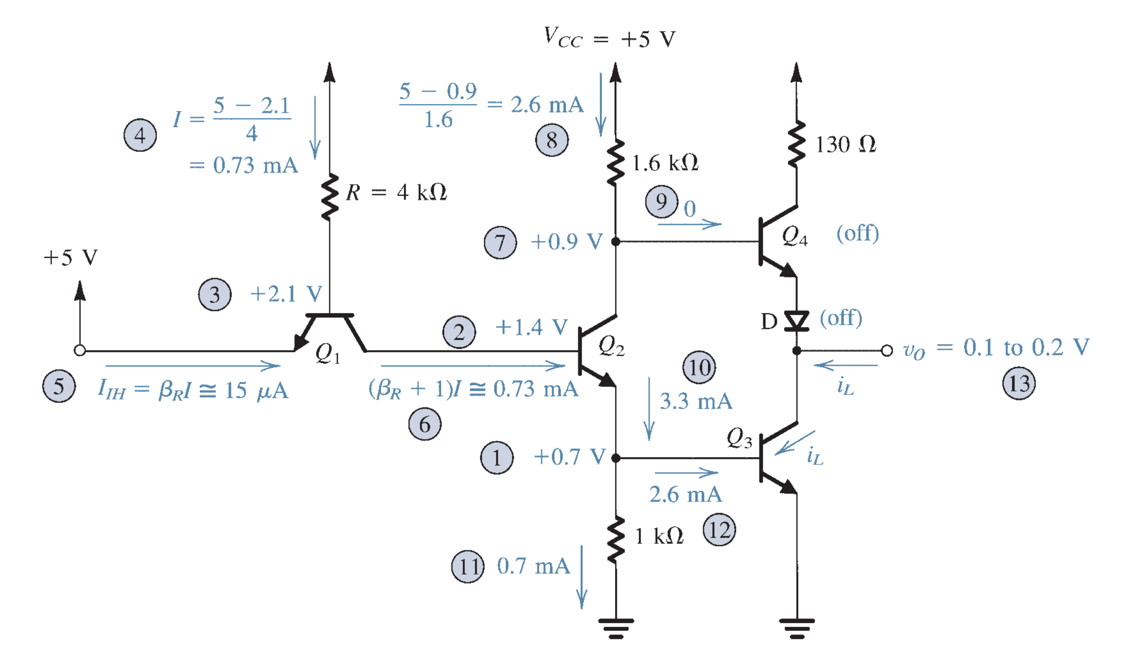


FIGURE 14.20 Analysis of the TTL gate with the input high. The circled numbers indicate the order of the analysis steps.

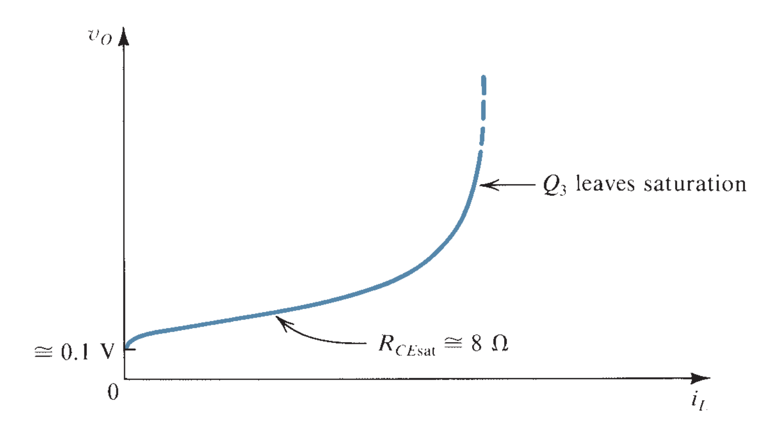


FIGURE 14.21 The v_O–i_L characteristic of the TTL gate when the output is low.

The saturated transistor Q_3 then establishes the low output voltage of the gate (V_{CEsat}) and provides a low impedance to ground.

In the low-output state the gate can sink a load current i_L, provided that the value of i_L does not exceed $\beta \times 2.6$ mA, which is the maximum collector current that Q_3 can sustain while remaining in saturation. Obviously the greater the value of i_L, the greater the output voltage will be. To maintain the logic-0 level below a certain specified limit, a corresponding limit has to be placed on the load current i_L. As will be seen shortly, it is this limit that determines the maximum fan-out of the TTL gate.

Figure 14.21 shows a sketch of the output voltage v_O versus the load current i_L of the TTL gate when the output is low. This is simply the v_{CE}–i_C characteristic curve of Q_3 measured with a base current of 2.6 mA. Note that at $i_L = 0$, v_O is the offset voltage, which is about 100 mV.

EXERCISE

14.5 Assume that the saturation portion of the v_O–i_L characteristic shown in Fig. 14.21 can be approximated by a straight line (of slope = 8 Ω) that intersects the v_O axis at 0.1 V. Find the maximum load current that the gate is allowed to sink if the logic-0 level is specified to be ≤0.3 V.

Ans. 25 mA

Analysis When the Input Is Low

Consider next the operation of the TTL gate when the input is at the logic-0 level (≃0.2 V). The analysis is illustrated in Fig. 14.22, from which we see that the base–emitter junction of Q_1 will be forward-biased and the base voltage will be approximately +0.9 V. Thus the current I can be found to be approximately 1 mA. Since 0.9 V is insufficient to forward-bias the series combination of the collector–base junction of Q_1 and the base–emitter junction of Q_2 (at least 1.2 V would be required), the latter will be off. Therefore the collector current of Q_1 will be almost zero and Q_1 will be saturated, with $V_{CEsat} \simeq 0.1$ V. Thus the base of Q_2 will be at approximately +0.3 V, which is indeed insufficient to turn Q_2 on.

The gate input current in the low state, called **input-low current** I_{IL}, is approximately equal to the current I ($\simeq 1$ mA) and flows out of the emitter of Q_1. If the TTL gate is driven

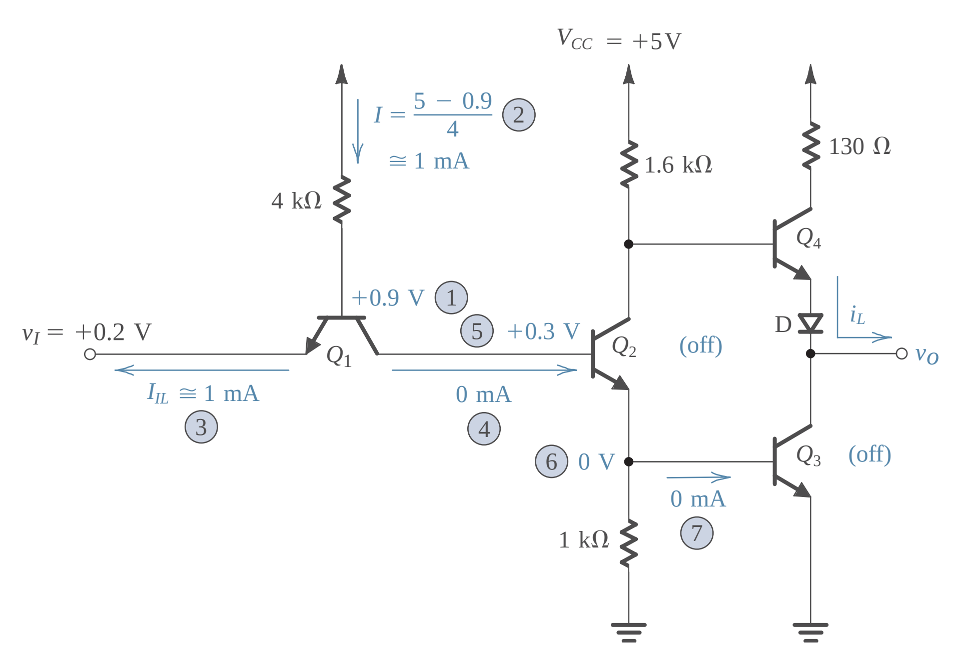


FIGURE 14.22 Analysis of the TTL gate when the input is low. The circled numbers indicate the order of the analysis steps.

by another TTL gate, the output transistor Q_3 of the driving gate should sink this current I_{IL}. Since the output current that a TTL gate can sink is limited to a certain maximum value, the maximum fan-out of the gate is directly determined by the value of I_{IL}.

EXERCISES

14.6 Consider the TTL gate analyzed in Exercise 14.5. Find its maximum allowable fan-out using the value of I_{IL} calculated above.

Ans. 25

14.7 Use Eq. (4.114) to find V_{CEsat} of transistor Q_1 when the input of the gate is low (0.2 V). Assume that $\beta_F = 50$ and $\beta_R = 0.02$.

Ans. 98 mV

Let us continue with our analysis of the TTL gate. When the input is low, we see that both Q_2 and Q_3 will be off. Transistor Q_4 will be on and will supply (source) the load current i_L. Depending on the value of i_L, Q_4 will be either in the active mode or in the saturation mode.

With the gate output terminal open, the current i_L will be very small (mostly leakage) and the two junctions (base–emitter junction of Q_4 and diode D) will be barely conducting. Assuming that each junction has a 0.65-V drop and neglecting the voltage drop across the 1.6-kΩ resistance, we find that the output voltage will be

$$v_O \simeq 5 - 0.65 - 0.65 = 3.7 \text{ V}$$

As i_L is increased, Q_4 and D conduct more heavily, but for a range of i_L, Q_4 remains in the active mode, and v_O is given by

$$v_O = V_{CC} - \frac{i_L}{\beta + 1} \times 1.6 \text{ k}\Omega - V_{BE4} - V_D \tag{14.4}$$

If we keep increasing i_L, a value will be reached at which Q_4 saturates. Then the output voltage becomes determined by the 130-Ω resistance according to the approximate relationship

$$v_O \simeq V_{CC} - i_L \times 130 - V_{CEsat}(Q_4) - V_D \tag{14.5}$$

Function of the 130-Ω Resistance

At this point the reason for including the 130-Ω resistance should be evident: It is simply to limit the current that flows through Q_4, especially in the event that the output terminal is accidentally short-circuited to ground. This resistance also limits the supply current in another circumstance, namely, when Q_4 turns on while Q_3 is still in saturation. To see how this occurs, consider the case where the gate input was high and then is suddenly brought down to the low level. Transistor Q_2 will turn off relatively fast because of the availability of a large reverse current supplied to its base terminal by the collector of Q_1. On the other hand, the base of Q_3 will have to discharge through the 1-kΩ resistance, and thus Q_3 will take some time to turn off. Meanwhile Q_4 will turn on, and a large current pulse will flow through the series combination of Q_4 and Q_3. Part of this current will serve the useful purpose of charging up any load capacitance to the logic-1 level. The magnitude of the current pulse will be limited by the 130-Ω resistance to about 30 mA.

The occurrence of these current pulses of short duration (called **current spikes**) raises another important issue. The current spikes have to be supplied by the V_{CC} source and, because of its finite source resistance, will result in voltage spikes (or "glitches") superimposed on V_{CC}. These voltage spikes could be coupled to other gates and flip-flops in the digital system and thus might produce false switching in other parts of the system. This effect, which might loosely be called **crosstalk,** is a problem in TTL systems. To reduce the size of the voltage spikes, capacitors (called bypass capacitors) should be connected between the supply rail and ground at frequent locations. These capacitors lower the impedance of the supply-voltage source and hence reduce the magnitude of the voltage spikes. Alternatively, one can think of the bypass capacitors as supplying the impulsive current spikes.

EXERCISES

14.8 Assuming that Q_4 has $\beta = 50$ and that at the verge of saturation $V_{CEsat} = 0.3$ V, find the value of i_L at which Q_4 saturates.

Ans. 4.16 mA

14.9 Assuming that at a current of 1 mA the voltage drops across the emitter–base junction of Q_4 and the diode D are each 0.7 V, find v_O when $i_L = 1$ mA and 10 mA. (Note the result of the previous exercise.)

Ans. 3.6 V; 2.7 V

14.10 Find the maximum current that can be sourced by a TTL gate while the output high level (V_{OH}) remains greater than the minimum guaranteed value of 2.4 V.

Ans. 12.3 mA; or, more accurately, taking the base current of Q_4 into account, 13.05 mA



14.4 CHARACTERISTICS OF STANDARD TTL

Because of its historical popularity and continued importance, TTL will be studied further in this and the next sections. In this section we shall consider some of the important characteristics of standard TTL gates. Special improved forms of TTL will be dealt with in Section 14.5.

Transfer Characteristic

Figure 14.23 shows the TTL gate together with a sketch of its voltage transfer characteristic drawn in a piecewise-linear fashion. The actual characteristic is, of course, a smooth curve. We shall now explain the transfer characteristic and calculate the various break-points and slopes. It will be assumed that the output terminal of the gate is open.

Segment *AB* is obtained when transistor Q_1 is saturated, Q_2 and Q_3 are off, and Q_4 and D are on. The output voltage is approximately two diode drops below V_{CC}. At point *B* the phase splitter (Q_2) begins to turn on because the voltage at its base reaches 0.6 V (0.5 V + V_{CEsat} of Q_1).

Over segment *BC*, transistor Q_1 remains saturated, but more and more of its base current *I* gets diverted to its base–collector junction and into the base of Q_2, which operates as a linear amplifier. Transistor Q_4 and diode D remain on, with Q_4 acting as an emitter follower. Meanwhile the voltage at the base of Q_3, although increasing, remains insufficient to turn Q_3 on (less than 0.6 V).

Let us now find the slope of segment *BC* of the transfer characteristic. Let the input v_I increase by an increment Δv_I. This increment appears at the collector of Q_1, since the saturated Q_1 behaves (approximately) as a three-terminal short circuit as far as signals are

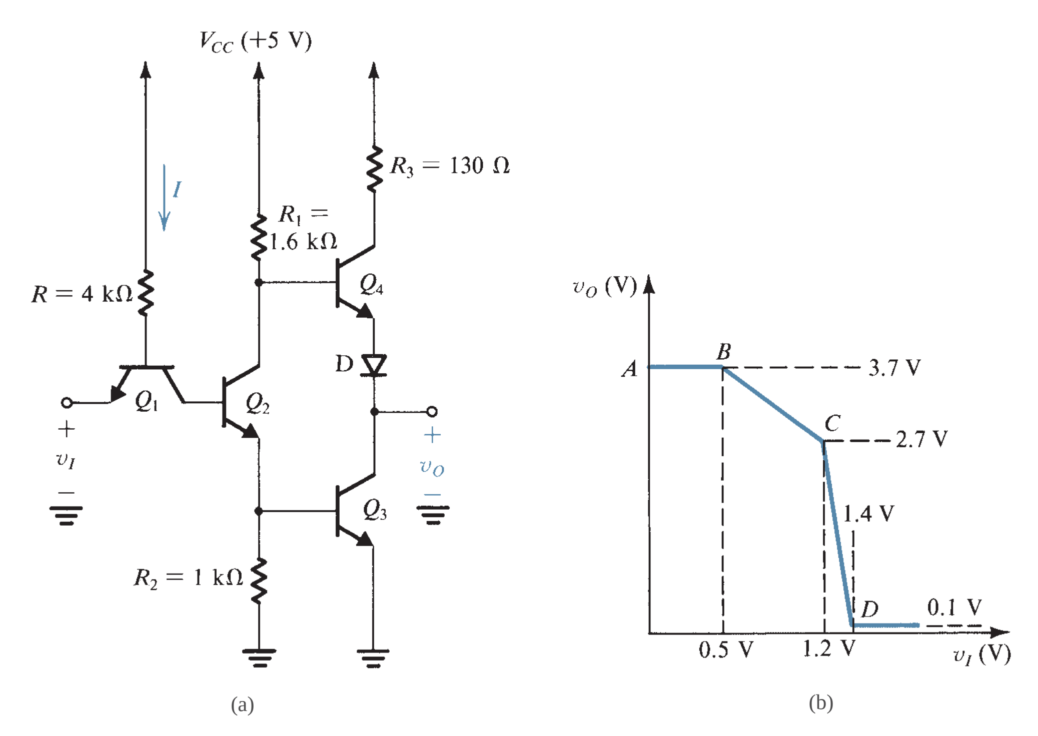


FIGURE 14.23 The TTL gate and its voltage transfer characteristic.

concerned. Thus at the base of Q_2 we have a signal Δv_I. Neglecting the loading of emitter follower Q_4 on the collector of Q_2, we can find the gain of the phase splitter from

$$\frac{v_{c2}}{v_{b2}} = \frac{-\alpha_2 R_1}{r_{e2} + R_2} \tag{14.6}$$

The value of r_{e2} will obviously depend on the current in Q_2. This current will range from zero (as Q_2 begins to turn on) to the value that results in a voltage of about 0.6 V at the emitter of Q_2 (the base of Q_3). This value is about 0.6 mA and corresponds to point C on the transfer characteristic. Assuming an average current in Q_2 of 0.3 mA, we obtain $r_{e2} \simeq 83\ \Omega$. For $\alpha = 0.98$, Eq. (14.6) results in a gain value of 1.45. Since the gain of the output follower Q_4 is close to unity, the overall gain of the gate, which is the slope of the BC segment, is about –1.45.

As already implied, breakpoint C is determined by Q_3 starting to conduct. The corresponding input voltage can be found from

$$\begin{aligned} v_I(C) &= V_{BE3} + V_{BE2} - V_{CE\text{sat}}(Q_1) \\ &= 0.6 + 0.7 - 0.1 = 1.2 \text{ V} \end{aligned}$$

At this point the emitter current of Q_2 is approximately 0.6 mA. The collector current of Q_2 is also approximately 0.6 mA; neglecting the base current of Q_4, the voltage at the collector of Q_2 is

$$v_{C2}(C) = 5 - 0.6 \times 1.6 \cong 4 \text{ V}$$

Thus Q_2 is still in the active mode. The corresponding output voltage is

$$v_O(C) = 4 - 0.65 - 0.65 = 2.7 \text{ V}$$

As v_I is increased past the value of $v_I(C) = 1.2$ V, Q_3 begins to conduct and operates in the active mode. Meanwhile, Q_1 remains saturated, and Q_2 and Q_4 remain in the active mode. The circuit behaves as an amplifier until Q_2 and Q_3 saturate and Q_4 cuts off. This occurs at point D on the transfer characteristic, which corresponds to an input voltage $v_I(D)$ obtained from

$$\begin{aligned} v_I(D) &= V_{BE3} + V_{BE2} + V_{BC1} - V_{BE1} \\ &= 0.7 + 0.7 + 0.7 - 0.7 = 1.4 \text{ V} \end{aligned}$$

Note that we have in effect assumed that at point D transistor Q_1 is still saturated, but with $V_{CE\text{sat}} \simeq 0$. To see how this comes about, note that from point B on, more and more of the base current of Q_1 is diverted to its base–collector junction. Thus while the drop across the base–collector junction increases, that across the base–emitter junction decreases. At point D these drops become almost equal. For $v_I > v_I(D)$ the base–emitter junction of Q_1 cuts off; thus Q_1 leaves saturation and enters the inverse active mode.

Calculation of gain over the segment CD is a relatively complicated task. This is due to the fact that there are two paths from input to output: one through Q_3 and one through Q_4. A simple but gross approximation for the gain of this segment can be obtained from the coordinates of points C and D in Fig. 14.23(b), as follows:

$$\begin{aligned} \text{Gain} &= -\frac{v_O(C) - v_O(D)}{v_I(D) - v_I(C)} \\ &= -\frac{2.7 - 0.1}{1.4 - 1.2} = -13 \text{ V/V} \end{aligned}$$

From the transfer curve of Fig. 14.23(b) we can determine the critical points and the noise margins as follows: $V_{OH} = 3.7$ V; V_{IL} is somewhere in the range of 0.5 V to 1.2 V, and thus a conservative estimate would be 0.5 V; $V_{OL} = 0.1$ V; $V_{IH} = 1.4$ V; $NM_H = V_{OH} - V_{IH} = 2.3$ V; and $NM_L = V_{IL} - V_{OL} = 0.4$ V. It should be noted that these values are computed assuming that the gate is not loaded and without taking into account power-supply or temperature variations.

EXERCISE

14.11 Taking into account the fact that the voltage across a forward-biased *pn* junction changes by about –2 mV/°C, find the coordinates of points *A*, *B*, *C*, and *D* of the gate transfer characteristic at –55°C and at +125°C. Assume that the characteristic in Fig. 14.23(b) applies at 25°C, and neglect the small temperature coefficient of V_{CEsat}.

Ans. At –55°C: (0, 3.38), (0.66, 3.38), (1.52, 2.16), (1.72, 0.1); at +125°C: (0, 4.1), (0.3, 4.1), (0.8, 3.46), (1.0, 0.1)

Manufacturers' Specifications

Manufacturers of TTL usually provide curves for the gate transfer characteristic, the input *i–v* characteristic, and the output *i–v* characteristic, measured at the limits of the specified operating temperature range. In addition, guaranteed values are usually given for the parameters V_{OL}, V_{OH}, V_{IL}, and V_{IH}. For standard TTL (known as the 74 series) these values are $V_{OL} = 0.4$ V, $V_{OH} = 2.4$ V, $V_{IL} = 0.8$ V, and $V_{IH} = 2$ V. These limit values are guaranteed for a specified tolerance in power-supply voltage and for a maximum fan-out of 10. From our discussion in Section 14.3 we know that the maximum fan-out is determined by the maximum current that Q_3 can sink while remaining in saturation and while maintaining a saturation voltage lower than a guaranteed maximum ($V_{OL} = 0.4$ V). Calculations performed in Section 14.3 indicate the possibility of a maximum fan-out of 20 to 30. Thus the figure specified by the manufacturer is appropriately conservative.

The parameters V_{OL}, V_{OH}, V_{IL}, and V_{IH} can be used to compute the noise margins as follows:

$$NM_H = V_{OH} - V_{IH} = 0.4 \text{ V}$$

$$NM_L = V_{IL} - V_{OL} = 0.4 \text{ V}$$

EXERCISES

14.12 In Section 14.3 we found that when the gate input is high, the base current of Q_3 is approximately 2.6 mA. Assume that this value applies at 25°C and that at this temperature $V_{BE} \simeq 0.7$ V. Taking into account the –2-mV/°C temperature coefficient of V_{BE} and neglecting all other changes, find the base current of Q_3 at –55°C and at +125°C.

Ans. 2.2 mA; 3 mA

14.13 Figure E14.13 shows sketches of the i_L–v_O characteristics of a TTL gate when the output is low. Use these characteristics together with the results of Exercise 14.12 to calculate the value of β of transistor Q_3 at –55°C, +25°C, and +125°C.

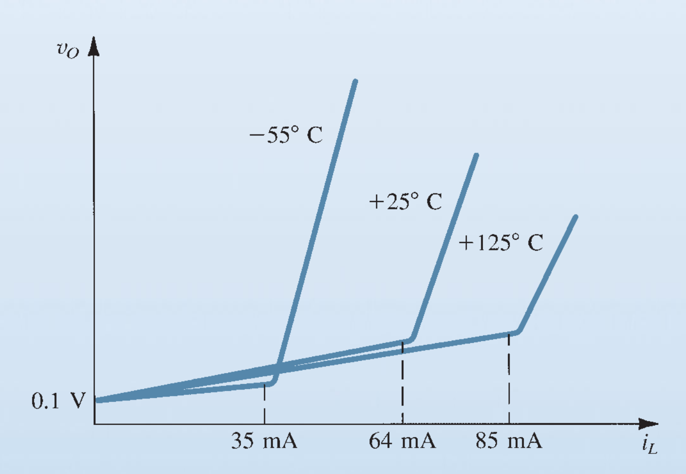


FIGURE E14.13

Ans. 16; 25; 28

Propagation Delay

The propagation delay of TTL gates is defined conventionally as the time between the 1.5-V points of corresponding edges of the input and output waveforms. For standard TTL (also known as *medium-speed* TTL) t_P is typically about 10 ns.

As far as power dissipation is concerned it can be shown (see Exercise 14.14) that when the gate output is high the gate dissipates 5 mW, and when the output is low the dissipation is 16.7 mW. Thus the average dissipation is 11 mW, resulting in a delay-power product of about 100 pJ.

EXERCISE

14.14 Calculate the value of the supply current (I_{CC}), and hence the power dissipated in the TTL gate, when the output terminal is open and the input is (a) low at 0.2 V (see Fig. 14.22) and (b) high at +5 V (see Fig. 14.20).

Ans. (a) 1 mA, 5 mW; (b) 3.33 mA, 16.7 mW

Dynamic Power Dissipation

In Section 14.3 the occurrence of supply current spikes was explained. These spikes give rise to additional power drain from the V_{CC} supply. This **dynamic power** is also dissipated in the gate circuit. It can be evaluated by multiplying the average current due to the spikes by V_{CC}, as illustrated by the solution of Exercise 14.15.

EXERCISE

14.15 Consider a TTL gate that is switched on and off at the rate of 1 MHz. Assume that each time the gate is turned off (that is, the output goes high) a supply-current pulse of 30-mA amplitude and 2-ns width occurs. Also assume that no current spike occurs when the gate is turned on. Calculate the average supply current due to the spikes, and the dynamic power dissipation.

Ans. 60 μA; 0.3 mW

The TTL NAND Gate

Figure 14.24 shows the basic TTL gate. Its most important feature is the **multiemitter transistor** Q_1 used at the input. Figure 14.25 shows the structure of the multiemitter transistor.

It can be easily verified that the gate of Fig. 14.24 performs the NAND function. The output will be high if one (or both) of the inputs is (are) low. The output will be low in only

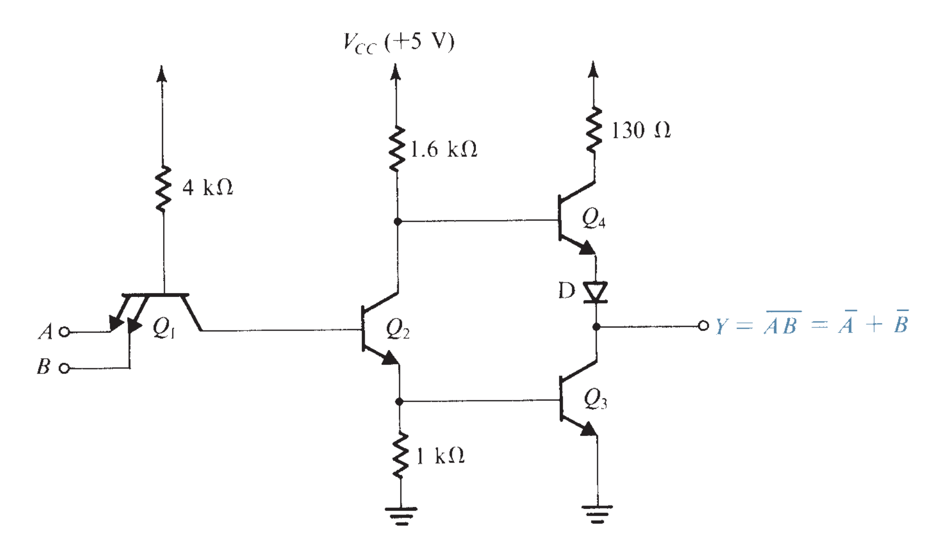


FIGURE 14.24 The TTL NAND gate.

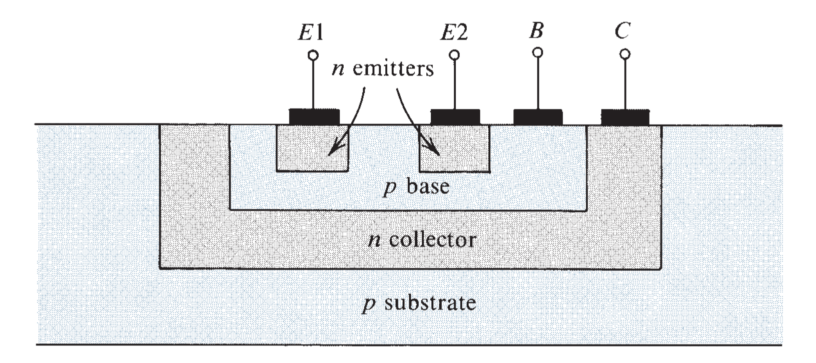


FIGURE 14.25 Structure of the multiemitter transistor Q_1.

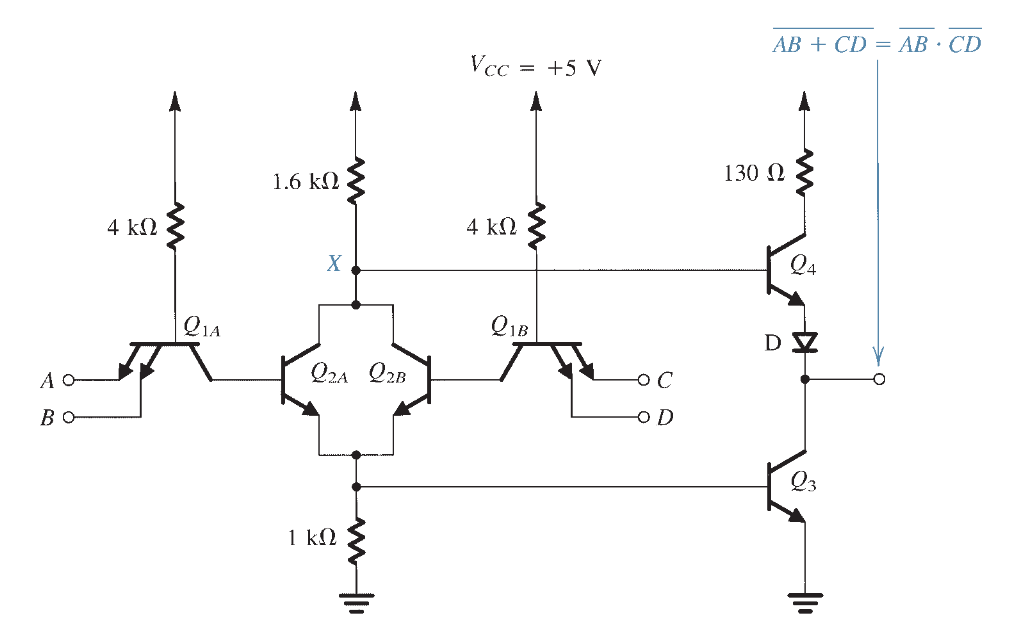


FIGURE 14.26 A TTL AND-OR-INVERT gate.

one case: when both inputs are high. Extension to more than two inputs is straightforward and is achieved by diffusing additional emitter regions.

Although theoretically an unused input terminal may be left open-circuited, this is generally not a good practice. An open-circuit input terminal acts as an "antenna" that "picks up" interfering signals and thus could cause erroneous gate switching. An unused input terminal should therefore be connected to the positive power supply *through a resistance* (of, say, 1 kΩ). In this way the corresponding base–emitter junction of Q_1 will be reverse-biased and thus will have no effect on the operation of the gate. The series resistance is included in order to limit the current in case of breakdown of the base–emitter junction due to transients on the power supply.

Other TTL Logic Circuits

On a TTL MSI chip there are many cases in which logic functions are implemented using "stripped-down" versions of the basic TTL gate. As an example we show in Fig. 14.26 the TTL implementation of the AND-OR-INVERT function. As shown, the phase-splitter transistors of two gates are connected in parallel, and a single output stage is used. The reader is urged to verify that the logic function realized is as indicated.

At this point it should be noted that the totem-pole output stage of TTL does *not* allow connecting the output terminals of two gates to realize the AND function of their outputs (known as the wired-AND connection). To see the reason for this, consider two gates whose outputs are connected together, and let one gate have a high output and the other have a low output. Current will flow from Q_4 of the first gate through Q_3 of the second gate. The current value will fortunately be limited by the 130-Ω resistance. Obviously, however, no useful logic function is realized by this connection.

The lack of wired-AND capability is a drawback of TTL. Nevertheless, the problem is solved in a number of ways, including doing the paralleling at the phase-splitter stage, as illustrated in Fig. 14.26. Another solution consists of deleting the emitter-follower transistor

altogether. The result is an output stage consisting solely of the common-emitter transistor Q_3 without even a collector resistance. Obviously, one can connect the outputs of such gates together to a common collector resistance and achieve a wired-AND capability. TTL gates of this type are known as **open-collector TTL.** The obvious disadvantage is the slow rise time of the output waveform.

Another useful variant of TTL is the **tristate** output arrangement explored in Exercise 14.16.

EXERCISE

14.16 The circuit shown in Fig. E14.16 is called tristate TTL. Verify that when the terminal labeled $\overline{\text{Third state}}$ is high, the gate functions normally and that when this terminal is low, both transistors Q_3 and Q_4 cut off and the output of the gate is an open circuit. The latter state is the third state, or the high-output-impedance state.

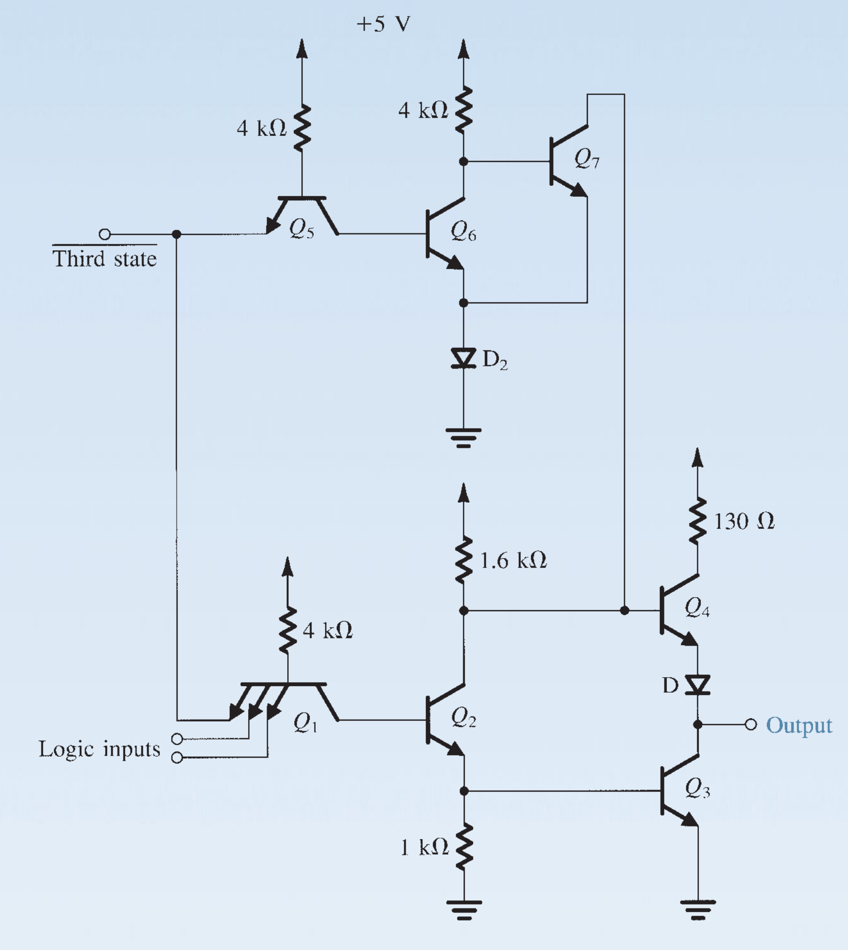


FIGURE E14.16

Tristate TTL enables the connection of a number of TTL gates to a common output line (or *bus*). At any particular time the signal on the bus will be determined by the one TTL gate

that is *enabled* (by raising its third-state input terminal). All other gates will be in the third state and thus will have no control of the bus.

14.5 TTL FAMILIES WITH IMPROVED PERFORMANCE

The standard TTL circuits studied in the two previous sections were introduced in the mid-1960s. Since then, several improved versions have been developed. In this section we shall discuss some of these improved TTL subfamilies. As will be seen the improvements are in two directions: increasing speed and reducing power dissipation.

The speed of the standard TTL gate of Fig. 14.24 is limited by two mechanisms: first, transistors Q_1, Q_2, and Q_3 saturate, and hence we have to contend with their finite storage time. Although Q_2 is discharged reasonably quickly because of the active mode of operation of Q_1, as already explained, this is not true for Q_3, whose base charge has to leak out through the 1-kΩ resistance in its base circuit. Second, the resistances in the circuit, together with the various transistor and wiring capacitances, form relatively long time constants, which contribute to lengthening the gate delay.

It follows that there are two approaches to speeding up the operation of TTL. The first is to prevent transistor saturation and the second is to reduce the values of all resistances. Both approaches are utilized in the Schottky TTL circuit family.

Schottky TTL

In Schottky TTL, transistors are prevented from saturation by connecting a low-voltage-drop diode between base and collector, as shown in Fig. 14.27. These diodes, formed as a metal-to-semiconductor junction, are called Schottky diodes and have a forward voltage drop of about 0.5 V. We have briefly discussed Schottky diodes in Section 3.9. Schottky diodes[4] are easily fabricated and do not increase chip area. In fact, the Schottky TTL fabrication process has been designed to yield transistors with smaller areas and thus higher β and f_T than those produced by the standard TTL process. Figure 14.27 also shows the symbol used to represent the combination of a transistor and a Schottky diode, referred to as a Schottky transistor.

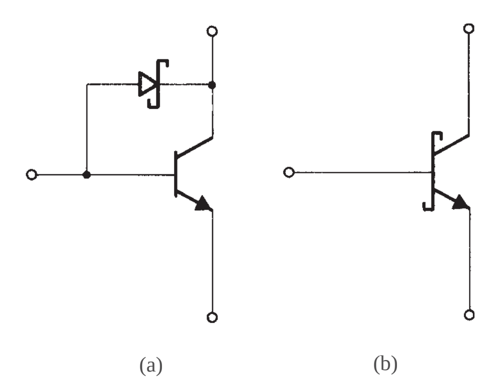


FIGURE 14.27 **(a)** A transistor with a Schottky diode clamp. **(b)** Circuit symbol for the connection in (a), known as a Schottky transistor.

[4] Note that silicon Schottky diodes exhibit voltage drops of about 0.5 V, whereas GaAs Schottky diodes (Section 5.12) exhibit voltage drops of about 0.7 V.

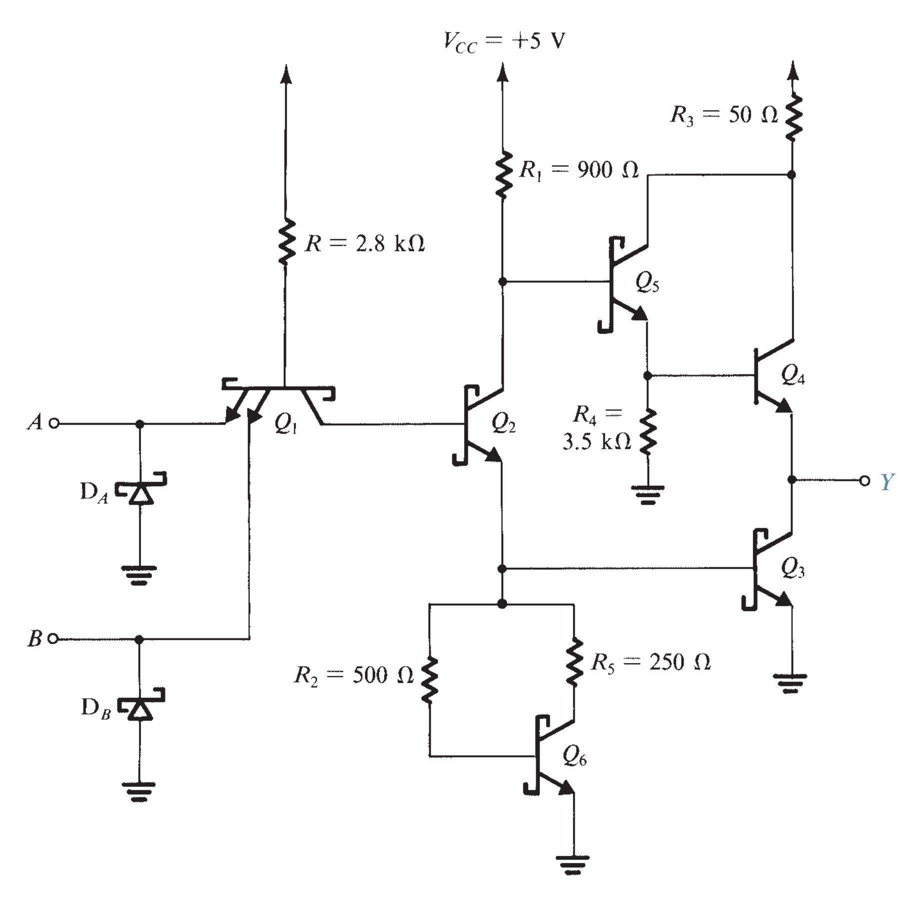


FIGURE 14.28 A Schottky TTL (known as STTL) NAND gate.

The Schottky transistor does not saturate since some of its base current drive is shunted away from the base by the Schottky diode. The latter then conducts and clamps the base-collector junction voltage to about 0.5 V. This voltage is smaller than the value required to forward-bias the collector-base junction of these small-sized transistors. In fact, a Schottky transistor begins to conduct when its v_{BE} is about 0.7 V, and the transistor is fully conducting when v_{BE} is about 0.8 V. With the Schottky diode clamping v_{BC} to about 0.5 V, the collector-to-emitter voltage is about 0.3 V and the transistor is still operating in the active mode. By avoiding saturation, the Schottky-clamped transistor exhibits a very short turnoff time.

Figure 14.28 shows a Schottky-clamped, or simply Schottky, TTL NAND gate. Comparing this circuit to that of standard TTL shown in Fig. 14.24 reveals a number of variations. First and foremost, Schottky clamps have been added to all transistors except Q_4. As will be seen shortly, transistor Q_4 can never saturate and thus does not need a Schottky clamp. Second, all resistances have been reduced to almost half the values used in the standard circuit. These two changes result in a much shorter gate delay. The reduction in resistance values, however, increases the gate power dissipation (by a factor of about 2).

The Schottky TTL gate features three other circuit techniques that further improve performance. These are as follows:

1. The diode D needed to prevent Q_4 from conducting when the gate output is low is replaced by transistor Q_5, which together with Q_4 forms a Darlington pair. This

Darlington stage provides increased current gain and hence increased current-sourcing capability. This, together with the lower output resistance of the gate (in the output-high state), yields a reduction in the time required to charge the load capacitance to the high level. Note that transistor Q_4 never saturates because

$$V_{CE4} = V_{CE5} + V_{BE4}$$
$$\simeq 0.3 + 0.8 = 1.1 \text{ V}$$

2. Input clamping diodes, D_A and D_B, are included.[5] These diodes conduct only when the input voltages go below ground level. This could happen due to "ringing" on the wires connecting the input of the gate to the output of another gate. Ringing occurs because such connecting wires behave as *transmission lines* that are not properly terminated. Without the clamping diodes, ringing can cause the input voltage to transiently go sufficiently negative so as to cause the substrate-to-collector junction of Q_1 (see Fig. 14.25) to become forward-biased. This in turn would result in improper gate operation. Also, ringing can cause the input voltage to go sufficiently positive to result in false gate switching. The input diodes clamp the negative excursions of the input ringing signal (to –0.5 V). Their conduction also provides a power loss in the transmission line, which results in *damping* of the ringing waveform and thus a reduction in its positive-going part.
3. The resistance between the base of Q_3 and ground has been replaced by a nonlinear resistance realized by transistor Q_6, and two resistors, R_2 and R_5. This nonlinear resistance is known as an **active pull-down,** in analogy to the active pull-up provided by the emitter-follower part of the totem-pole output stage. This feature is an ingenious one, and we shall discuss it in more detail.

Active Pull-Down Figure 14.29 shows a sketch of the *i–v* characteristic of the nonlinear network composed of Q_6 and its two associated resistors, R_2 and R_5. Also shown for comparison

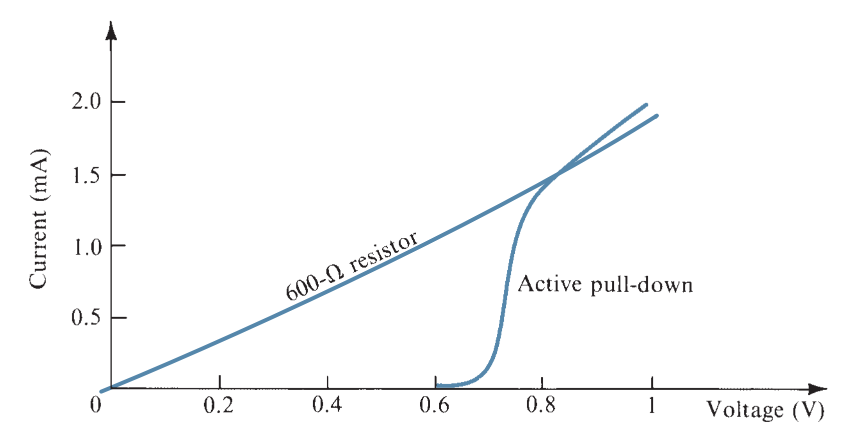


FIGURE 14.29 Comparison of the *i–v* characteristic of the active pull-down with that of a 600-Ω resistor.

[5] Some standard TTL circuits also include input clamping diodes. Note, however, that in Schottky TTL, the input clamping diodes are of the Schottky type and function at a lower voltage and have no minority-carrier stored charge.

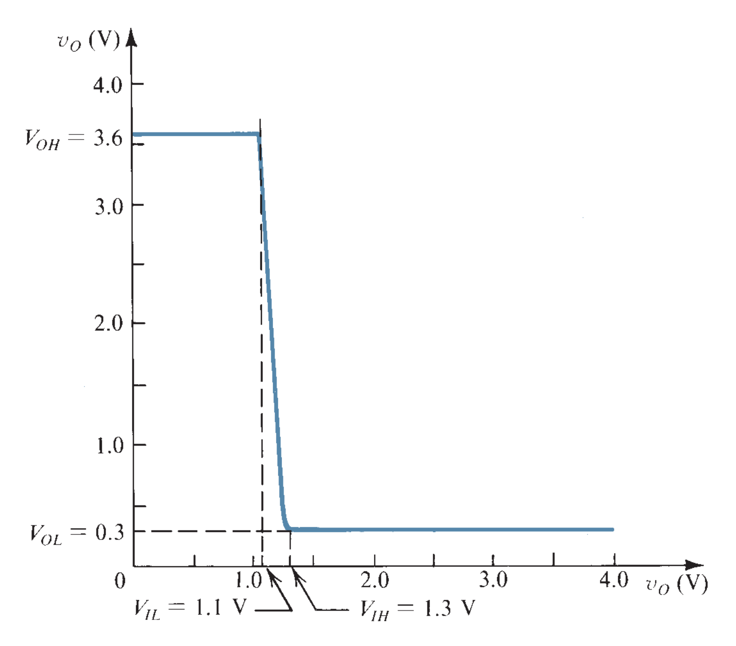


FIGURE 14.30 Voltage transfer characteristic of the Schottky TTL gate.

is the linear *i*–*v* characteristic of a resistor that would be connected between the base of Q_3 and ground if the active pull-down were not used.

The first characteristic to note of the active pull-down is that it conducts negligible current (and thus behaves as a high resistance) until the voltage across it reaches a V_{BE} drop. Thus the gain of the phase splitter (as a linear amplifier) will remain negligibly small until a V_{BE} drop develops between its emitter and ground, which is the onset of conduction of Q_3. In other words, Q_2, Q_3, and Q_6 will turn on almost simultaneously. Thus segment *BC* of the transfer characteristic (see Fig. 14.23b) will be absent and the gate transfer characteristic will become much sharper, as shown in Fig. 14.30. Since the active pull-down circuit causes the transfer characteristic to become "squarer," it is also known as a "squaring circuit." The result is an increase in noise margins.

The active pull-down circuit also speeds up the turn-on and turnoff of Q_3. To see how this comes about observe that the active pull-down draws a negligible current over a good part of its characteristic. Thus the current supplied by the phase splitter will initially be diverted into the base of Q_3, thereby speeding up the turn-on of Q_3. On the other hand, during turnoff the active pull-down transistor Q_6 will be operating in the active mode. Its large collector current will flow through the base of Q_3 in the reverse direction, thus quickly discharging the base-emitter capacitance of Q_3 and turning it off fast.

Finally, the function of the 250-Ω collector resistance of Q_6 should be noted; without it the base-emitter voltage of Q_3 would be clamped to V_{CE6}, which is about 0.3 V.

EXERCISES

14.17 Show that the values of V_{OH}, V_{IL}, V_{IH}, and V_{OL} of the Schottky TTL gate are as given in Fig. 14.30. Assume that the gate output current is small and that a Schottky transistor conducts at $v_{BE} = 0.7$ V and is fully conducting at $v_{BE} = 0.8$ V. (*Hint:* V_{IL} is the input voltage at which Q_2 and Q_3 begin to turn on; V_{IH} is the value at which Q_2 and Q_3 are fully conducting.)

14.18 Calculate the input current of a Schottky TTL gate with the input voltage low (at 0.3 V).

Ans. 1.4 mA

14.19 For the Schottky TTL gate calculate the current drawn from the power supply with the input low at 0.3 V (remember to include the current in the 3.5-kΩ resistor) and with the input high at 3.6 V. Hence calculate the gate power dissipation in both states and the average power dissipation.

Ans. 2.6 mA; 5.4 mA; 13 mW; 27 mW; 20 mW

Performance Characteristics Schottky TTL (known as Series 74S) is specified to have the following worst-case parameters:

$$V_{OH} = 2.7 \text{ V} \qquad V_{OL} = 0.5 \text{ V}$$
$$V_{IH} = 2.0 \text{ V} \qquad V_{IL} = 0.8 \text{ V}$$
$$t_P = 3 \text{ ns} \qquad P_D = 20 \text{ mW}$$

Note that the delay-power product is 60 pJ, as compared with about 100 pJ for standard TTL.

EXERCISE

14.20 Use the values given above to calculate the noise margins NM_H and NM_L for a Schottky TTL gate.

Ans. 0.7 V; 0.3 V

Low-Power Schottky TTL

Although Schottky TTL achieves a very low propagation delay, the gate power dissipation is rather high. This limits the number of gates that can be included per package. The need therefore arose for a modified version that achieves a lower gate dissipation, possibly at the expense of an increase in gate delay. This is obtained in the low-power Schottky (LS) TTL subfamily, represented by the two-input NAND gate shown in Fig. 14.31.

The low-power Schottky circuit of Fig. 14.31 differs from that of the regular Schottky TTL of Fig. 14.28 in several ways: most important, note that the resistances used are about ten times larger, with the result that the power dissipation is only about a tenth of that of the Schottky circuit (2 mW versus 20 mW). However, as expected, the use of larger resistances is accompanied by a reduction in the speed of operation. To help compensate for this speed reduction, some circuit design innovations are employed: First, the input multiemitter transistor has been eliminated in favor of Schottky diodes, which occupy a smaller silicon area and hence have smaller parasitic capacitances. In this regard it should be recalled that the main advantage of the input multiemitter transistor is that it rapidly removes the charge stored in the base of Q_2. In Schottky-clamped circuits, however, transistor Q_2 does not saturate, eliminating this aspect of the need for Q_1.

Second, two Schottky diodes, D_3 and D_4, have been added to the output stage to help speed up the turnoff of Q_4 and the turn-on of Q_3 and hence the transition of the output from high to low. Specifically, as the gate input is raised and Q_2 begins to turn on, some of its collector current will flow through diode D_3. This current constitutes a reverse base current for Q_4 and thus aids in the rapid turnoff of Q_4. Simultaneously, the emitter current of Q_2 will be supplied to the base of Q_3, thus causing it to turn on faster. Some of the collector current of Q_2 will also flow through D_4. This current will help discharge the gate load capacitance and thus shorten the transition time from high to low. Both D_3 and D_4 will be cut off under static conditions.

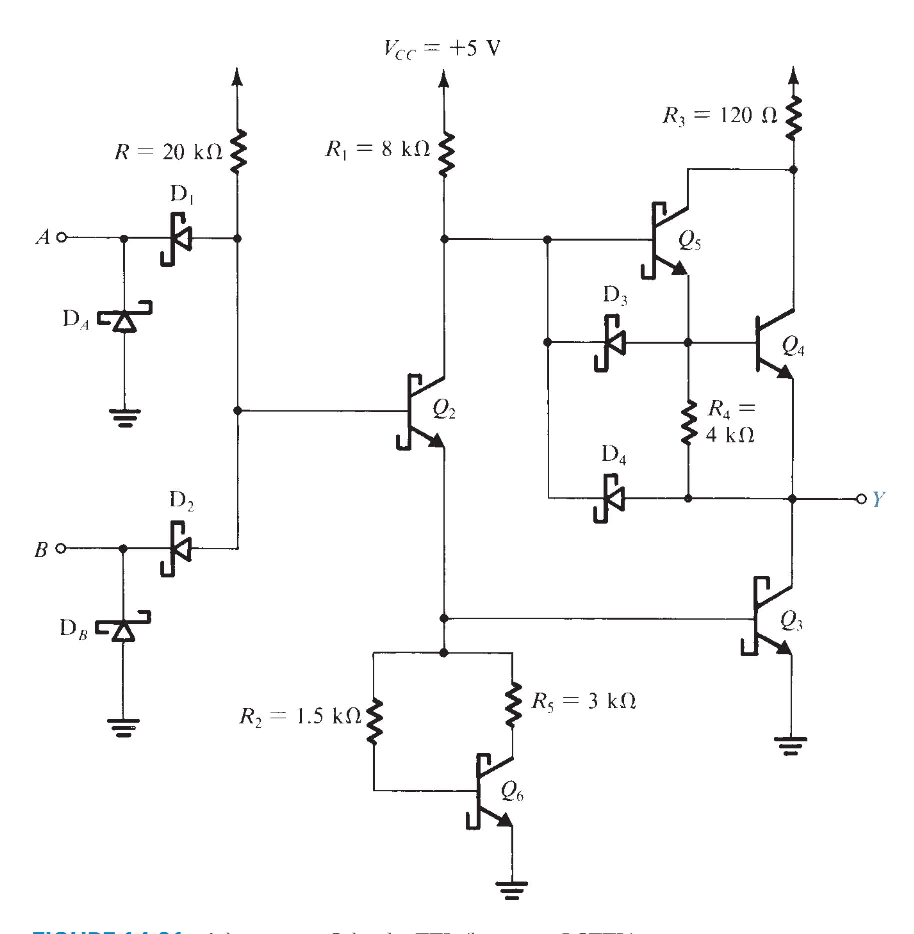


FIGURE 14.31 A low-power Schottky TTL (known as LSTTL) gate.

Finally, observe that the other terminal of the emitter resistor of Q_5 is now connected to the gate output. It follows that when the gate output is high, the output current will be first supplied by Q_5 through R_4. Transistor Q_4 will turn on only when a 0.7-V drop develops across R_4, and thereafter it sources additional load current. However, when the gate is supplying a small load current, Q_4 will be off and the output voltage will be approximately.

$$V_{OH} = V_{CC} - V_{BE5}$$

which is higher than the value obtained in the regular Schottky TTL gate.

EXERCISES

14.21 The voltage transfer characteristic of the low-power Schottky TTL gate of Fig. 14.31 has the same shape as that of the regular Schottky TTL gate, shown in Fig. 14.30. Calculate the values of V_{OH}, V_{IL}, V_{IH}, and V_{OL}. Assume that a Schottky transistor conducts at $v_{BE} = 0.7$ V and is fully conducting at $v_{BE} =$ 0.8 V, and that a Schottky diode has a 0.5-V drop. Also assume that the gate output current is very small.

Ans. 4.3 V; 0.9 V; 1.1 V; 0.3 V

14.22 For the low-power Schottky TTL gate, using the specifications given in Exercise 14.21, calculate the supply current in both states. Hence calculate the average power dissipation.

Ans. 0.21 mA; 0.66 mA; 2 mW

TABLE 14.1 Performance Comparison of TTL Families

	Standard TTL (Series 74)	Schottky TTL (Series 74S)	Low-Power Schottky TTL (Series 74LS)	Advanced Schottky TTL (Series 74AS)	Advanced Low-Power Schottky TTL (Series 74ALS)
t_P, ns	10	3	10	1.5	4
P_D, mW	10	20	2	20	1
DP, pJ	100	60	20	30	4

Although the guaranteed voltage levels and noise margins of low-power Schottky TTL (known as Series 74LS) are similar to those of the regular Schottky TTL, the gate delay and power dissipation are

$$t_P = 10 \text{ ns} \qquad P_D = 2 \text{ mW}$$

Thus although the power dissipation has been reduced by a factor of 10, the delay has been increased by only a factor of 3. The result is a delay-power product of only 20 pJ.

Further-Improved TTL Families

There are other TTL families with further improved characteristics. Of particular interest is the advanced Schottky (Series 74AS) and the advanced low-power Schottky (Series 74ALS). We shall not discuss the circuit details of these families here. Table 14.1 provides a comparison of the TTL subfamilies based on gate delay and power dissipation.

Observe that advanced low-power Schottky offers a very small delay-power product. In conclusion, we note that at the present time TTL, although no longer as popular as it once was, is still utilized in systems assembled using SSI and MSI chips. This is particularly the case for the advanced Schottky family whose speed of operation is rivaled only by that achieved in emitter-coupled logic (ECL). TTL, however, is not used in VLSI design.

15.4 Emitter-Coupled Logic (ECL)

Emitter-coupled logic (ECL) is the fastest logic circuit family available for conventional logic-system design.[4] High speed is achieved by operating all bipolar transistors out of saturation, thus avoiding storage-time delays, and by keeping the logic signal swings relatively small (about 0.8 V or less), thus reducing the time required to charge and discharge the various load and parasitic capacitances. Saturation in ECL is avoided by using the BJT differential pair as a current switch.[5] The BJT differential pair was studied in Chapter 9, and we urge the reader to review the introduction given in Section 9.2 before proceeding with the study of ECL.

15.4.1 The Basic Principle

Emitter-coupled logic is based on the use of the current-steering switch introduced in Section 15.6. Such a switch can be most conveniently realized using the differential pair shown in Fig. 15.25. The pair is biased with a constant-current source I, and one side is connected to a reference voltage V_R. As shown in Section 9.2, the current I can be steered to either Q_1 or Q_2 under the control of the input signal v_I. Specifically, when v_I is greater than V_R by about $4V_T$ ($\simeq$100 mV), nearly all the current I is conducted by Q_1, and thus for $\alpha_1 \simeq 1$, $v_{O1} = V_{CC} - IR_C$. Simultaneously, the current through Q_2 will be nearly zero, and thus $v_{O2} = V_{CC}$. Conversely, when v_I is lower than V_R by about $4V_T$, most of the current I will flow through Q_2 and the current through Q_1 will be nearly zero. Thus $v_{O1} = V_{CC}$ and $v_{O2} = V_{CC} - IR_C$.

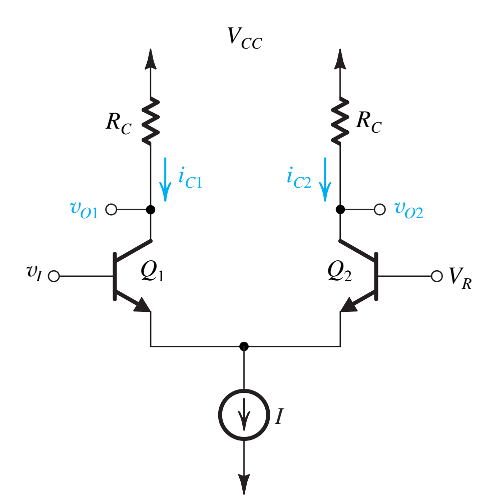


Figure 15.25 The basic element of ECL is the differential pair. Here, V_R is a reference voltage.

[4] Although higher speeds of operation can be obtained with gallium arsenide (GaAs) circuits, the latter are not available as off-the-shelf components for conventional digital system design. GaAs digital circuits are not covered in this book; however, a substantial amount of material on this subject can be found on the disc accompanying the book and on the website.

[5] This is in sharp contrast to the technique utilized in a nonsaturating variant of transistor-transistor logic (TTL) known as Schottky TTL. There, a Schottky diode is placed across the CBJ junction to shunt away some of the base current and, owing to the low voltage drop of the Schottky diode, the CBJ is prevented from becoming forward biased.

The preceding description suggests that as a logic element, the differential pair realizes an inversion function at v_{O1} and simultaneously provides the complementary output signal at v_{O2}. The output logic levels are $V_{OH} = V_{CC}$ and $V_{OL} = V_{CC} - IR_C$, and thus the output logic swing is IR_C. A number of additional remarks can be made concerning this circuit:

1. The differential nature of the circuit makes it less susceptible to picked-up noise. In particular, an interfering signal will tend to affect both sides of the differential pair similarly and thus will not result in current switching. This is the common-mode rejection property of the differential pair (see Section 9.2).
2. The current drawn from the power supply remains constant during switching. Thus, unlike CMOS (and TTL), no supply current spikes occur in ECL, eliminating an important source of noise in digital circuits. This is a definite advantage, especially since ECL is usually designed to operate with small signal swings and has correspondingly low noise margins.
3. The output signal levels are both referenced to V_{CC} and thus can be made particularly stable by operating the circuit with $V_{CC} = 0$: in other words, by utilizing a negative power supply and connecting the V_{CC} line to ground. In this case, $V_{OH} = 0$ and $V_{OL} = -IR_C$.
4. Some means must be provided to make the output signal levels compatible with those at the input so that one gate can drive another. As we shall see shortly, practical ECL gate circuits incorporate a level-shifting arrangement that serves to center the output signal levels on the value of V_R.
5. The availability of complementary outputs considerably simplifies logic design with ECL.

EXERCISE

15.11 For the circuit in Fig. 15.25, let $V_{CC} = 0$, $I = 4$ mA, $R_C = 220\,\Omega$, $V_R = -1.32$ V, and assume $\alpha \simeq 1$. Determine V_{OH} and V_{OL}. By how much should the output levels be shifted so that the values of V_{OH} and V_{OL} become centered on V_R? What will the shifted values of V_{OH} and V_{OL} be?
Ans. 0; –0.88 V; –0.88 V; –0.88 V, –1.76 V

15.4.2 ECL Families

Currently there are two popular forms of commercially available ECL—namely, ECL 10K and ECL 100K. The ECL 100K series features gate delays on the order of 0.75 ns and dissipates about 40 mW/gate, for a delay–power product of 30 pJ. Although its power dissipation is relatively high, the 100K series provides the shortest available gate delay in small- and medium-scale integrated circuit packages.

The ECL 10 K series is slightly slower; it features a gate propagation delay of 2 ns and a power dissipation of 25 mW for a delay–power product of 50 pJ. Although the value of *PDP* is higher than that obtained in the 100K series, the 10K series is easier to use. This is because the rise and fall times of the pulse signals are deliberately made longer, thus reducing signal coupling, or cross talk, between adjacent signal lines. ECL 10K has an “edge speed”

of about 3.5 ns, compared with the approximately 1 ns of ECL 100K. To give concreteness to our study of ECL, in the following we shall consider the popular ECL 10K in some detail. The same techniques, however, can be applied to other types of ECL.

In addition to its usage in SSI and MSI circuit packages, ECL is also employed in large-scale and VLSI applications. A variant of ECL known as **current-mode logic** (CML) is utilized in VLSI applications (see Treadway, 1989, and Wilson, 1990).

15.4.3 The Basic Gate Circuit

The basic gate circuit of the ECL 10K family is shown in Fig. 15.26. The circuit consists of three parts. The network composed of Q_1, D_1, D_2, R_1, R_2, and R_3 generates a reference voltage V_R whose value at room temperature is –1.32 V. As will be shown, the value of this reference voltage is made to change with temperature in a predetermined manner to keep the noise margins almost constant. Also, the reference voltage V_R is made relatively insensitive to variations in the power-supply voltage V_{EE}.

EXERCISE

15.12 Figure E15.12 shows the circuit that generates the reference voltage V_R. Assuming that the voltage drop across each of D_1, D_2, and the base–emitter junction of Q_1 is 0.75 V, calculate the value of V_R. Neglect the base current of Q_1.

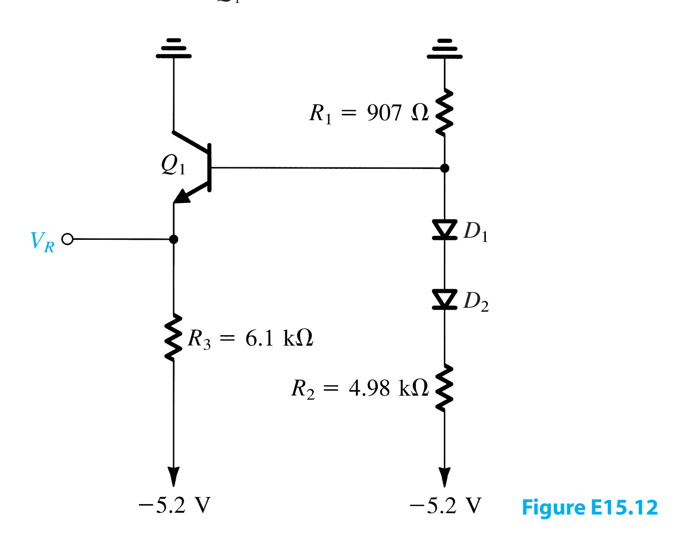


Figure E15.12

Ans. –1.32 V

The second part, and the heart of the gate, is the differential amplifier formed by Q_R and either Q_A or Q_B. This differential amplifier is biased not by a constant-current source, as was done in the circuit of Fig. 15.25, but with a resistance R_E connected to the negative supply $-V_{EE}$. Nevertheless, we will shortly show that the current in R_E remains approximately

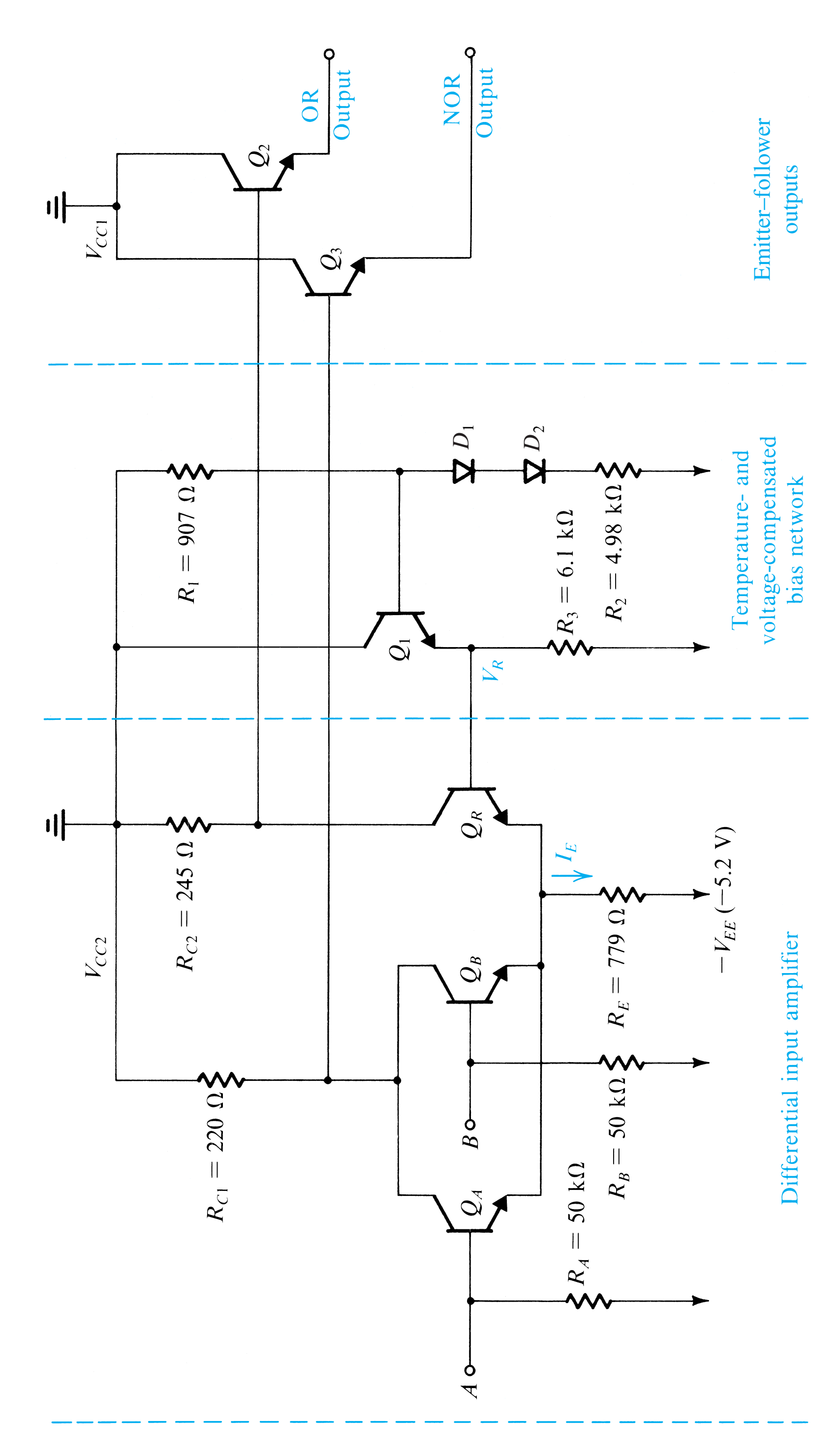


Figure 15.26 Basic circuit of the ECL 10K logic-gate family.

constant over the normal range of operation of the gate. One side of the differential amplifier consists of the reference transistor Q_R, whose base is connected to the reference voltage V_R. The other side consists of a number of transistors (two in the case shown), connected in parallel, with separated bases, each connected to a gate input. If the voltages applied to A and B are at the logic-0 level, which, as we will soon find out, is about 0.4 V below V_R, both Q_A and Q_B, will be off and the current I_E in R_E will flow through the reference transistor Q_R. The resulting voltage drop across R_{C2} will cause the collector voltage of Q_R to be low.

On the other hand, when the voltage applied to A or B is at the logic-1 level, which, as we will show shortly, is about 0.4 V above V_R, transistor Q_A or Q_B, or both, will be on and Q_R will be off. Thus the current I_E will flow through Q_A or Q_B, or both, and an almost equal current will flow through R_{C1}. The resulting voltage drop across R_{C1} will cause the collector voltage to drop. Meanwhile, since Q_R is off, its collector voltage rises. We thus see that the voltage at the collector of Q_R will be high if A or B, or both, is high, and thus at the collector of Q_R, the OR logic function, $A+B$, is realized. On the other hand, the common collector of Q_A and Q_B will be high only when A and B are simultaneously low. Thus at the common collector of Q_A and Q_B, the logic function $\overline{A}\,\overline{B}=\overline{A+B}$ is realized. We therefore conclude that the two-input gate of Fig. 15.26 realizes the OR function and its complement, the NOR function. The availability of complementary outputs is an important advantage of ECL; it simplifies logic design and avoids the use of additional inverters with associated time delay.

It should be noted that the resistance connecting each of the gate input terminals to the negative supply enables the user to leave an unused input terminal open: An open input terminal will be *pulled down* to the negative supply voltage, and its associated transistor will be off.

EXERCISE

15.13 With input terminals A and B in Fig. 15.26 left open, find the current I_E through R_E. Also find the voltages at the collector of Q_R and at the common collector of the input transistors Q_A and Q_B. Use $V_R=-1.32$ V, V_{BE} of $Q_R \simeq 0.75$ V, and assume that β of Q_R is very high.
Ans. 4 mA; –1 V; 0 V

The third part of the ECL gate circuit is composed of the two emitter followers, Q_2 and Q_3. The emitter followers do not have on-chip loads, since in many applications of high-speed logic circuits the gate output drives a transmission line terminated at the other end, as indicated in Fig. 15.27. (More on this later in Section 15.4.6.)

The emitter followers have two purposes: First, they shift the level of the output signals by one V_{BE} drop. Thus, using the results of Exercise 15.13, we see that the output levels become approximately –1.75 V and –0.75 V. These shifted levels are centered approximately around the reference voltage ($V_R=-1.32$ V), which means that one gate can drive another. This compatibility of logic levels at input and output is an essential requirement in the design of gate circuits.

The second function of the output emitter followers is to provide the gate with low output resistances and with the large output currents required for charging load capacitances. Since these large transient currents can cause spikes on the power-supply line, the collectors of

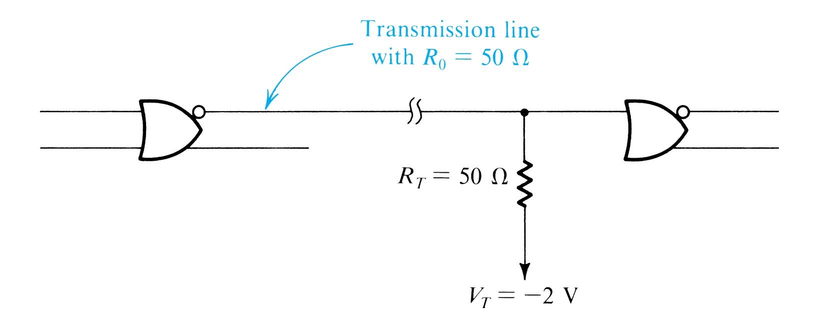


Figure 15.27 The proper way to connect high-speed logic gates such as ECL. Properly terminating the transmission line connecting the two gates eliminates the "ringing" that would otherwise corrupt the logic signals. (See Section 15.4.6.)

the emitter followers are connected to a power-supply terminal V_{CC1} separate from that of the differential amplifier and the reference-voltage circuit, V_{CC2}. Here we note that the supply current of the differential amplifier and the reference circuit remains almost constant. The use of separate power-supply terminals prevents the coupling of power-supply spikes from the output circuit to the gate circuit and thus lessens the likelihood of false gate switching. Both V_{CC1} and V_{CC2} are of course connected to the same system ground, external to the chip.

15.4.4 Voltage-Transfer Characteristics

Having provided a qualitative description of the operation of the ECL gate, we shall now derive its voltage-transfer characteristics. This will be done under the conditions that the outputs are terminated in the manner indicated in Fig. 15.27. Assuming that the *B* input is low and thus Q_B is off, the circuit simplifies to that shown in Fig. 15.28. We wish to analyze this circuit to determine v_{OR} versus v_I and v_{NOR} versus v_I (where $v_I \equiv v_A$).

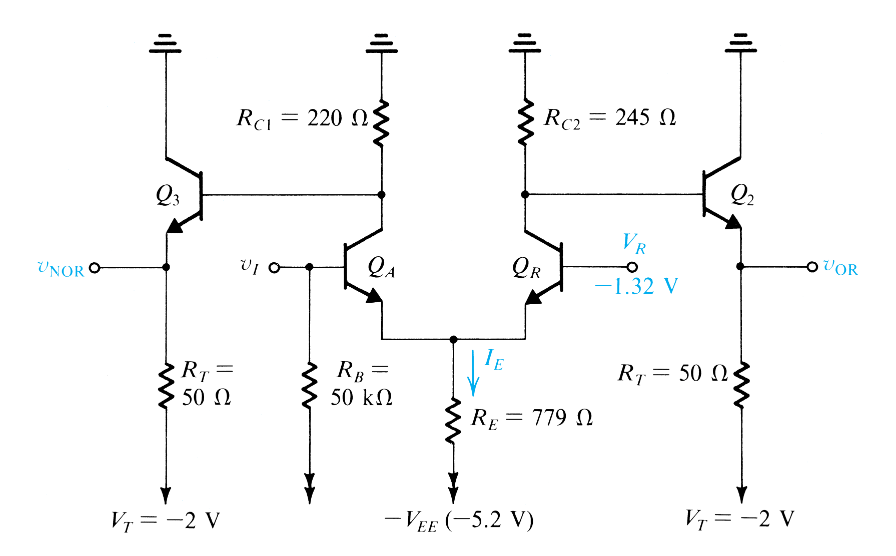


Figure 15.28 Simplified version of the ECL gate for the purpose of finding transfer characteristics.

In the analysis to follow we shall make use of the exponential $i_C - v_{BE}$ characteristic of the BJT. Since the BJTs used in ECL circuits have small areas (in order to have small capacitances and hence high f_T), their scale currents I_S are small. We will therefore assume that at an emitter current of 1 mA, an ECL transistor has a V_{BE} drop of 0.75 V.

The OR Transfer Curve Figure 15.29 is a sketch of the OR transfer characteristic, v_{OR} versus v_I, with the parameters V_{OL}, V_{OH}, V_{IL}, and V_{IH} indicated. However, to simplify the calculation of V_{IL} and V_{IH}, we shall use an alternative to the unity-gain definition. Specifically, we shall assume that at point x, transistor Q_A is conducting 1% of I_E while Q_R is conducting 99% of I_E. The reverse will be assumed for point y. Thus at point x we have

$$\frac{I_E|_{Q_R}}{I_E|_{Q_A}} = 99$$

Using the exponential $i_E - v_{BE}$ relationship, we obtain

$$V_{BE}|_{Q_R} - V_{BE}|_{Q_A} = V_T \ln 99 = 115 \text{ mV}$$

which gives

$$V_{IL} = -1.32 - 0.115 = -1.435 \text{ V}$$

Assuming Q_A and Q_R to be matched, we can write

$$V_{IH} - V_R = V_R - V_{IL}$$

which can be used to find V_{IH} as

$$V_{IH} = -1.205 \text{ V}$$

To obtain V_{OL}, we note that Q_A is off and Q_R carries the entire current I_E, given by

$$I_E = \frac{V_R - V_{BE}|_{Q_R} + V_{EE}}{R_E}$$

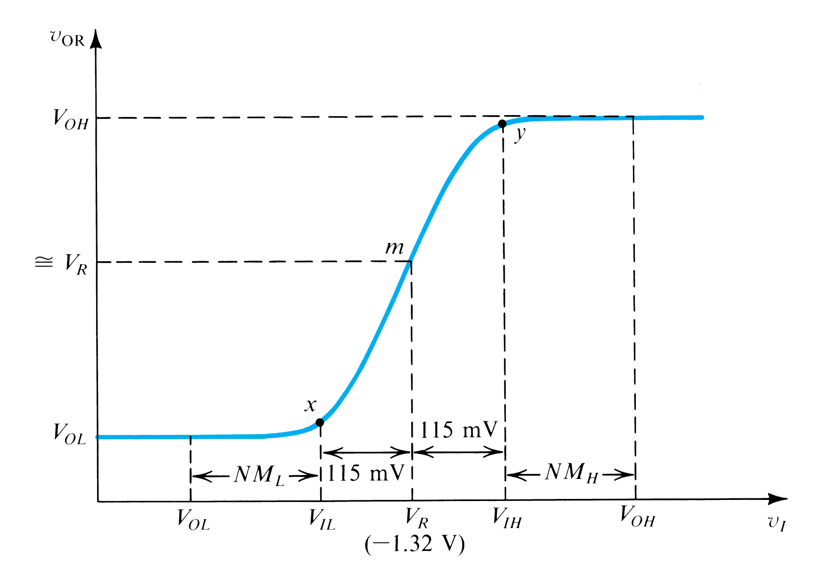


Figure 15.29 The OR transfer characteristic v_{OR} versus v_I, for the circuit in Fig. 15.28.

$$= \frac{-1.32 - 0.75 + 5.2}{0.779}$$

$$\simeq 4 \text{ mA}$$

(If we wish, we can iterate to determine a better estimate of $V_{BE}|_{Q_R}$ and hence of I_E.) Assuming that Q_R has a high β so that its $\alpha \simeq 1$, its collector current will be approximately 4 mA. If we neglect the base current of Q_2, we obtain for the collector voltage of Q_R

$$V_C|_{Q_R} \simeq -4 \times 0.245 = -0.98 \text{ V}$$

Thus a first approximation for the value of the output voltage V_{OL} is

$$V_{OL} = V_C|_{Q_R} - V_{BE}|_{Q_2}$$

$$\simeq -0.98 - 0.75 = -1.73 \text{ V}$$

We can use this value to find the emitter current of Q_2 and then iterate to determine a better estimate of its base–emitter voltage. The result is $V_{BE2} \simeq 0.79$ V and, correspondingly,

$$V_{OL} \simeq -1.77 \text{ V}$$

At this value of output voltage, Q_2 supplies a load current of about 4.6 mA.

To find the value of V_{OH} we assume that Q_R is completely cut off (because $v_I > V_{IH}$). Thus the circuit for determining V_{OH} simplifies to that in Fig. 15.30. Analysis of this circuit, assuming $\beta_2 = 100$, results in $V_{BE2} \simeq 0.83$ V, $I_{E2} = 22.4$ mA, and

$$V_{OH} \simeq -0.88 \text{ V}$$

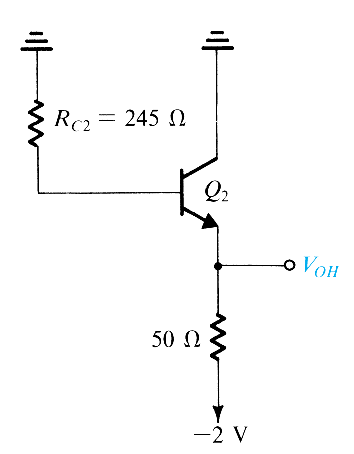


Figure 15.30 Circuit for determining V_{OH}.

EXERCISE

15.14 For the circuit in Fig. 15.28, determine the values of I_E obtained when $v_I = V_{IL}$, V_R, and V_{IH}. Also, find the value of v_{OR} corresponding to $v_I = V_R$. Assume that $v_{BE} = 0.75$ V at a current of 1 mA.
Ans. 3.97 mA; 4.00 mA; 4.12 mA; –1.31 V

Noise Margins The results of Exercise 15.14 indicate that the bias current I_E remains approximately constant. Also, the output voltage corresponding to $v_I = V_R$ is approximately equal to V_R. Notice further that this is also approximately the midpoint of the logic swing; specifically,

$$\frac{V_{OL} + V_{OH}}{2} = -1.325 \simeq V_R$$

Thus the output logic levels are centered around the midpoint of the input transition band. This is an ideal situation from the point of view of noise margins, and it is one of the reasons for selecting the rather arbitrary-looking numbers ($V_R = -1.32$ V and $V_{EE} = 5.2$ V) for reference and supply voltages.

The noise margins can now be evaluated as follows:

$$NM_H = V_{OH} - V_{IH} = -0.88 - (-1.205) = 0.325 \text{ V}$$

$$NM_L = V_{IL} - V_{OL} = -1.435 - (-1.77) = 0.335 \text{ V}$$

Note that these values are approximately equal.

The NOR Transfer Curve The NOR transfer characteristic, which is v_{NOR} versus v_I for the circuit in Fig. 15.28, is sketched in Fig. 15.31. The values of V_{IL} and V_{IH} are identical to those found earlier for the OR characteristic. To emphasize this, we have labeled the threshold points x and y, the same letters used in Fig. 15.29.

For $v_I < V_{IL}$, Q_A is off and the output voltage v_{NOR} can be found by analyzing the circuit composed of R_{C1}, Q_3, and its 50-Ω termination. Except that R_{C1} is slightly smaller than R_{C2}, this circuit is identical to that in Fig. 15.30. Thus the output voltage will be only slightly greater than the value V_{OH} found earlier. In the sketch of Fig. 15.31 we have assumed that the output voltage is approximately equal to V_{OH}.

For $v_I > V_{IH}$, Q_A is on and is conducting the entire bias current. The circuit then simplifies to that in Fig. 15.32. This circuit can be easily analyzed to obtain v_{NOR} versus v_I for the range

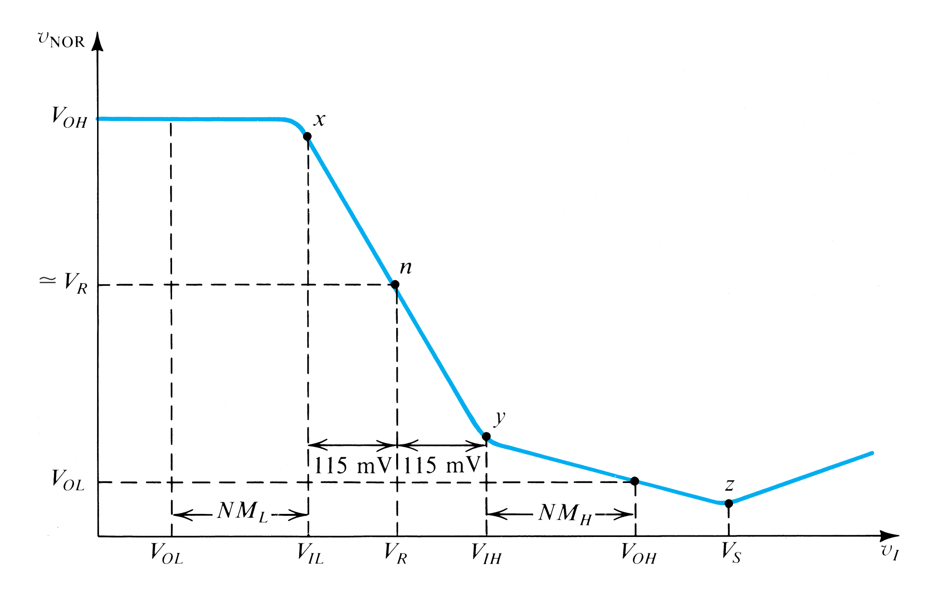


Figure 15.31 The NOR transfer characteristic, v_{NOR} versus v_I, for the circuit in Fig. 15.28.

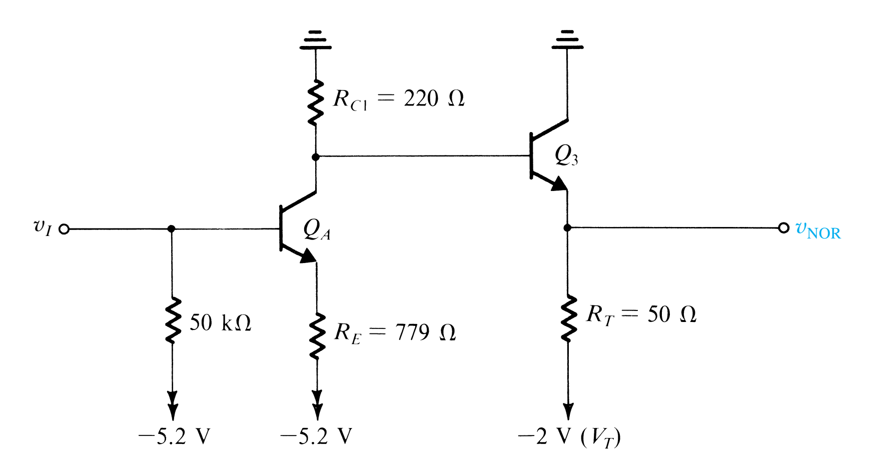


Figure 15.32 Circuit for finding v_{NOR} versus v_I for the range $v_I > V_{IH}$.

$v_I \geq V_{IH}$. A number of observations are in order. First, note that $v_I = V_{IH}$ results in an output voltage slightly higher than V_{OL}. This is because R_{C1} is smaller than R_{C2}. In fact, R_{C1} is chosen lower in value than R_{C2} so that with v_I equal to the normal logic-1 value (i.e., V_{OH}, which is approximately –0.88 V), the output will be equal to the V_{OL} value found earlier for the OR output.

Second, note that as v_I exceeds V_{IH}, transistor Q_A operates in the active mode and the circuit of Fig. 15.32 can be analyzed to find the gain of this amplifier, which is the slope of the segment *yz* of the transfer characteristic. At point *z*, transistor Q_A saturates. Further increments in v_I (beyond the point $v_I = V_S$) cause the collector voltage and hence v_{NOR} to increase. The slope of the segment of the transfer characteristic beyond point *z*, however, is not unity, but is about 0.5, because as Q_A is driven deeper into saturation, a portion of the increment in v_I appears as an increment in the base–collector forward-bias voltage. The reader is urged to solve Exercise 15.15, which is concerned with the details of the NOR transfer characteristic.

EXERCISE

15.15 Consider the circuit in Fig. 15.32. (a) For $v_I = V_{IH} = -1.205$ V, find v_{NOR}. (b) For $v_I = V_{OH} = -0.88$ V, find v_{NOR}. (c) Find the slope of the transfer characteristic at the point $v_I = V_{OH} = -0.88$ V. (d) Find the value of v_I at which Q_A saturates (i.e., V_S). Assume that $V_{BE} = 0.75$ V at a current of 1 mA, $V_{CE\text{sat}} \simeq 0.3$ V, and $\beta = 100$.

Ans. (a) –1.70 V; (b) –1.79 V; (c) –0.24 V/V; (d) –0.58 V

Manufacturers' Specifications ECL manufacturers supply gate transfer characteristics of the form shown in Figs. 15.29 and 15.31. A manufacturer usually provides such curves measured at a number of temperatures. In addition, at each relevant temperature, worst-case values for the parameters V_{IL}, V_{IH}, V_{OL}, and V_{OH} are given. These worst-case values are specified with the inevitable component tolerances taken into account. As an example, Motorola specifies that for MECL 10,000 at 25°C, the following worst-case

values apply[6]

$$V_{IL\max} = -1.475 \text{ V} \qquad V_{IH\min} = -1.105 \text{ V}$$
$$V_{OL\max} = -1.630 \text{ V} \qquad V_{OH\min} = -0.980 \text{ V}$$

These values can be used to determine worst-case noise margins,

$$NM_L = 0.155 \text{ V} \qquad NM_H = 0.125 \text{ V}$$

which are about half the *typical* values previously calculated.

For additional information on MECL specifications the interested reader is referred to the Motorola (1988, 1989) publications listed in the bibliography in Appendix I.

15.4.5 Fan-Out

When the input signal to an ECL gate is low (V_{OL}), the input current is equal to the current that flows in the 50-kΩ pull-down resistor. Thus,

$$I_{IL} = \frac{-1.77 + 5.2}{50} \simeq 69\,\mu\text{A}$$

When the input is high (V_{OH}), the input current is greater because of the base current of the input transistor. Thus, assuming a transistor β of 100, we obtain

$$I_{IH} = \frac{-0.88 + 5.2}{50} + \frac{4}{101} \simeq 126\ \mu\text{A}$$

Both these current values are quite small, which, coupled with the very small output resistance of the ECL gate, ensures that little degradation of logic-signal levels results from the input currents of fan-out gates. It follows that the fan-out of ECL gates is not limited by logic-level considerations but rather by the degradation of the circuit speed (rise and fall times). This latter effect is due to the capacitance that each fan-out gate presents to the driving gate (approximately 3 pF). Thus while the *dc fan-out* can be as high as 90 and thus does not represent a design problem, the *ac fan-out* is limited by considerations of circuit speed to 10 or so.

15.4.6 Speed of Operation and Signal Transmission

The speed of operation of a logic family is measured by the delay of its basic gate and by the rise and fall times of the output waveforms. Typical values of these parameters for ECL have already been given. Here we should note that because the output circuit is an emitter follower, the rise time of the output signal is shorter than its fall time, since on the rising edge of the output pulse, the emitter follower functions and provides the output current required to charge up the load and parasitic capacitances. On the other hand, as the signal at the base of the emitter follower falls, the emitter follower cuts off, and the load capacitance discharges through the combination of load and pull-down resistances.

To take full advantage of the very high speed of operation possible with ECL, special attention should be paid to the method of interconnecting the various logic gates in a system. To appreciate this point, we shall briefly discuss the problem of signal transmission.

ECL deals with signals whose rise times may be 1 ns or even less, the time it takes for light to travel only 30 cm or so. For such signals, a wire and its environment become a relatively complex circuit element along which signals propagate with finite speed (perhaps half the speed of light—i.e., 15 cm/ns). Unless special care is taken, energy that reaches the end

[6]MECL is the trade name used by Motorola (now Freescale Semiconductors) for its ECL.

of such a wire is not absorbed but rather returns as a *reflection* to the transmitting end, where (without special care) it may be re-reflected. The result of this process of reflection is what can be observed as **ringing**, a damped oscillatory excursion of the signal about its final value.

Unfortunately, ECL is particularly sensitive to ringing because the signal levels are so small. Thus it is important that transmission of signals be well controlled, and surplus energy absorbed, to prevent reflections. The accepted technique is to limit the nature of connecting wires in some way. One way is to insist that they be very short, where "short" is taken to mean with respect to the signal rise time. The reason for this is that if the wire connection is so short that reflections return while the input is still rising, the result becomes only a somewhat slowed and "bumpy" rising edge.

If, however, the reflection returns *after* the rising edge, it produces not simply a modification of the initiating edge but an *independent second event.* This is clearly bad! Thus the time taken for a signal to go from one end of a line and back is restricted to less than the rise time of the driving signal by some factor—say, 5. Thus for a signal with a 1-ns rise time and for propagation at the speed of light (30 cm/ns), a double path of only 0.2-ns equivalent length, or 6 cm, would be allowed, representing in the limit a wire only 3 cm from end to end.

Such is the restriction on ECL 100K. However, ECL 10K has an intentionally slower rise time of about 3.5 ns. Using the same rules, wires can accordingly be as long as about 10 cm for ECL 10K.

If greater lengths are needed, then transmission lines must be used. These are simply wires in a controlled environment in which the distance to a ground reference plane or a second wire is highly controlled. Thus they might simply be twisted pairs of wires, one of which is grounded, or parallel ribbon wires, every second of which is grounded, or so-called microstrip lines on a printed-circuit board. The latter are simply copper strips of controlled geometry on one side of a thin printed-circuit board, the other side of which consists of a grounded plane.

Such transmission lines have a *characteristic impedance*, R_0, that ranges from a few tens of ohms to hundreds of ohms. Signals propagate on such lines somewhat more slowly than the speed of light, perhaps half as fast. When a transmission line is terminated at its receiving end in a resistance equal to its characteristic impedance, R_0, all the energy sent on the line is absorbed at the receiving end, and no reflections occur (since the termination acts as a limitless length of transmission line). Thus, signal integrity is maintained. Such transmission lines are said to be *properly terminated.* A properly terminated line appears at its sending end as a resistor of value R_0. The followers of ECL 10K with their open emitters and low output resistances (specified to be 7 Ω maximum) are ideally suited for driving transmission lines. ECL is also good as a line receiver. The simple gate with its high (50-kΩ) pull-down input resistor represents a very high resistance to the line. Thus a few such gates can be connected to a terminated line with little difficulty. Both these ideas are represented in Fig. 15.27.

15.4.7 Power Dissipation

Because of the differential-amplifier nature of ECL, the gate current remains approximately constant and is simply steered from one side of the gate to the other depending on the input logic signals. Thus, the supply current and hence the gate power dissipation of unterminated ECL remain relatively constant independent of the logic state of the gate. It follows that no voltage spikes are introduced on the supply line. Such spikes can be a dangerous source of noise in a digital system. It follows that in ECL the need for supply-line bypassing[7] is not as great as in, say, TTL. This is another advantage of ECL.

[7]Achieved by connecting capacitances to ground at frequent intervals along the power-supply line on a printed-circuit board.

At this juncture we should reiterate a point we made earlier, namely, that although an ECL gate would operate with $V_{EE} = 0$ and $V_{CC} = +5.2$ V, the selection of $V_{EE} = -5.2$ V and $V_{CC} = 0$ V is recommended, because in the circuit, all signal levels are referenced to V_{CC}, and ground is certainly an excellent reference.

EXERCISE

15.16 For the ECL gate in Fig. 15.26, calculate an approximate value for the power dissipated in the circuit under the condition that all inputs are low and that the emitters of the output followers are left open. Assume that the reference circuit supplies four identical gates, and hence only a quarter of the power dissipated in the reference circuit should be attributed to a single gate.
Ans. 22.4 mW

15.4.8 Thermal Effects

In our analysis of the ECL gate of Fig. 15.26, we found that at room temperature the reference voltage V_R is –1.32 V. We have also shown that the midpoint of the output logic swing is approximately equal to this voltage, which is an ideal situation in that it results in equal high and low noise margins. In Example 15.4, we shall derive expressions for the temperature coefficients of the reference voltage and of the output low and high voltages. In this way, it will be shown that the midpoint of the output logic swing varies with temperature at the same rate as the reference voltage. As a result, although the magnitudes of the high and low noise margins change with temperature, their values remain equal. This is an added advantage of ECL and provides a demonstration of the high degree of design optimization of this gate circuit.

Example 15.4

We wish to determine the temperature coefficient of the reference voltage V_R and of the midpoint between V_{OL} and V_{OH}.

Solution

To determine the temperature coefficient of V_R, consider the circuit in Fig. E15.12 and assume that the temperature changes by $+1$°C. Denoting the temperature coefficient of the diode and transistor voltage drops by δ, where $\delta \simeq -2$ mV/°C, we obtain the equivalent circuit shown in Fig. 15.33. In the latter circuit, the changes in device voltage drops are considered as signals, and hence the power supply is shown as a signal ground.

In the circuit of Fig. 15.33 we have two signal generators, and we wish to analyze the circuit to determine ΔV_R, the change in V_R. We shall do so using the principle of superposition.[8] Consider first the branch R_1, D_1, D_2, 2δ, and R_2, and neglect the signal base current of Q_1. The voltage signal at the base of Q_1 can be easily obtained from

[8]Although the circuit contains diodes and a transistor, which are nonlinear elements, we can use superposition because we are dealing with small changes in voltages and currents, and thus the diodes and the transistor are replaced by their linear small-signal models.

Example 15.4 *continued*

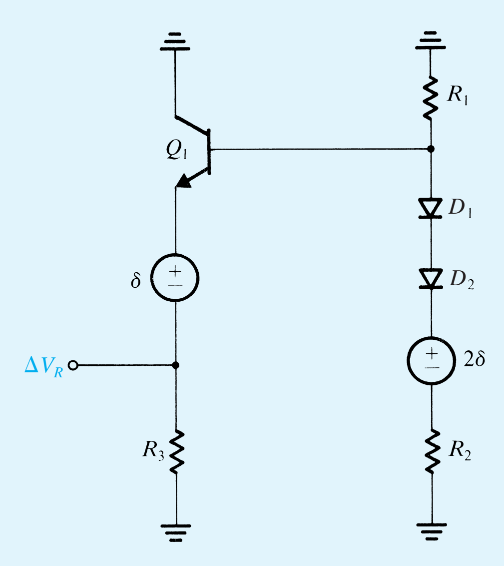


Figure 15.33 Equivalent circuit for determining the temperature coefficient of the reference voltage V_R.

$$v_{b1} = \frac{2\delta \times R_1}{R_1 + r_{d1} + r_{d2} + R_2}$$

where r_{d1} and r_{d2} denote the incremental resistances of diodes D_1 and D_2, respectively. The dc bias current through D_1 and D_2 is approximately 0.64 mA, and thus $r_{d1} = r_{d2} = 39.5\,\Omega$. Hence $v_{b1} \simeq 0.3\delta$. Since the gain of the emitter follower Q_1 is approximately unity, it follows that the component of ΔV_R due to the generator 2δ is approximately equal to v_{b1}, that is, $\Delta V_{R1} = 0.3\delta$.

Consider next the component of ΔV_R due to the generator δ. Reflection into the emitter circuit of the total resistance of the base circuit, $[R_1 \| (r_{d1} + r_{d2} + R_2)]$, by dividing it by $\beta + 1$ (with $\beta \simeq 100$) results in the following component of ΔV_R:

$$\Delta V_{R2} = -\frac{\delta \times R_3}{[R_B/(\beta+1)] + r_{e1} + R_3}$$

Here R_B denotes the total resistance in the base circuit, and r_{e1} denotes the emitter resistance of Q_1 ($\simeq 40\,\Omega$). This calculation yields $\Delta V_{R2} \simeq -\delta$. Adding this value to that due to the generator 2δ gives $\Delta V_{\mathrm{R}} \simeq -0.7\delta$. Thus for $\delta = -2$ mV/°C the temperature coefficient of V_R is +1.4 mV/°C.

We next consider the determination of the temperature coefficient of V_{OL}. The circuit on which to perform this analysis is shown in Fig. 15.34. Here we have three generators whose contributions can be considered separately and the resulting components of ΔV_{OL} summed. The result is

$$\begin{aligned}\Delta V_{OL} = \Delta V_R \frac{-R_{C2}}{r_{eR} + R_E} \frac{R_T}{R_T + r_{e2}} \\ - \delta \frac{-R_{C2}}{r_{eR} + R_E} \frac{R_T}{R_T + r_{e2}} \\ - \delta \frac{R_T}{R_T + r_{e2} + R_{C2}/(\beta+1)}\end{aligned}$$

Example 15.4 *continued*

Substituting the values given and those obtained throughout the analysis of this section, we find

$$\Delta V_{OL} \simeq -0.43\delta$$

The circuit for determining the temperature coefficient of V_{OH} is shown in Fig. 15.35, from which we obtain

$$\Delta V_{OH} = -\delta \frac{R_T}{R_T + r_{e2} + R_{C2}/(\beta+1)} = -0.93\delta$$

We now can obtain the variation of the midpoint of the logic swing as

$$\frac{\Delta V_{OL} + \Delta V_{OH}}{2} = -0.68\delta$$

which is approximately equal to that of the reference voltage $V_R(-0.7\delta)$.

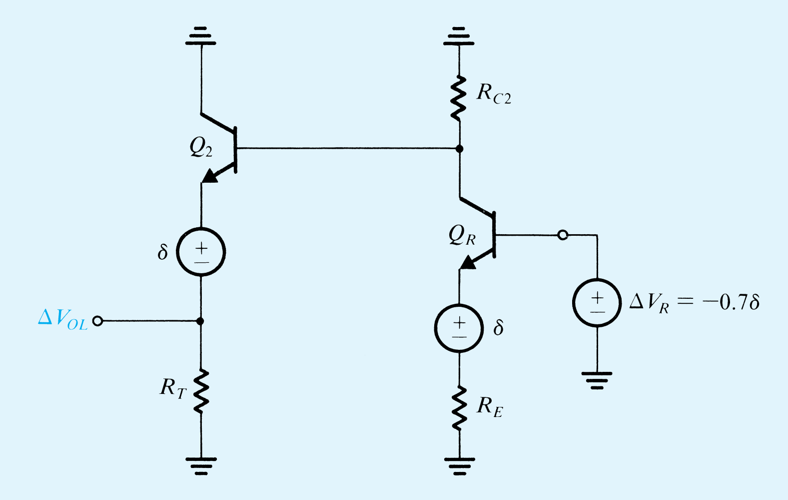


Figure 15.34 Equivalent circuit for determining the temperature coefficient of V_{OL}.

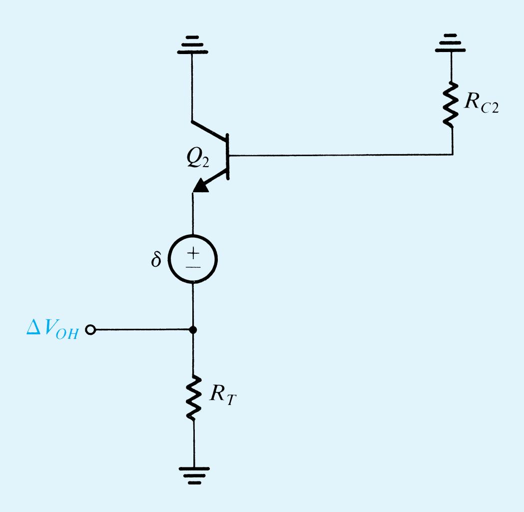


Figure 15.35 Equivalent circuit for determining the temperature coefficient of V_{OH}.

15.4.9 The Wired-OR Capability

The emitter–follower output stage of the ECL family allows an additional level of logic to be performed at very low cost by simply wiring the outputs of several gates in parallel. This is illustrated in Fig. 15.36, where the outputs of two gates are wired together. Note that the base–emitter diodes of the output followers realize an OR function: This **wired-OR** connection can be used to provide gates with high fan-in as well as to increase the flexibility of ECL in logic design.

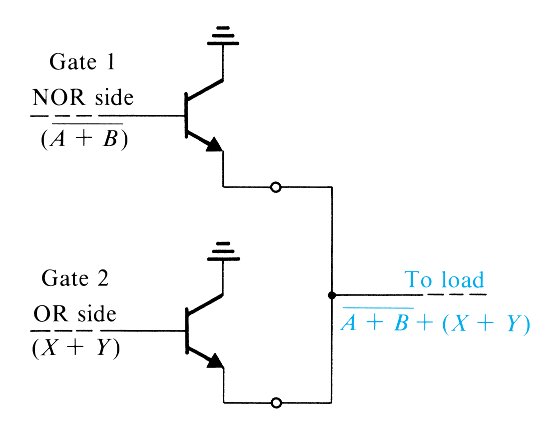


Figure 15.36 The wired-OR capability of ECL.

15.4.10 Final Remarks

We have chosen to study ECL by focusing on a commercially available circuit family. As has been demonstrated, a great deal of design optimization has been applied to create a very-high-performance family of SSI and MSI logic circuits. As already mentioned, ECL and some of its variants are also used in VLSI circuit design. Applications include very-high-speed processors such as those used in supercomputers, as well as high-speed and high-frequency communication systems. When employed in VLSI design, current–source biasing is almost always utilized. Further, a variety of circuit configurations are employed (see Rabaey, 1996).

A 15.5 BiCMOS Digital Circuits

In this section, we provide an introduction to a VLSI circuit technology that is becoming increasingly popular, BiCMOS. As its name implies, BiCMOS technology combines *bi*polar and *CMOS* circuits on one IC chip. The aim is to combine the low-power, high-input impedance and wide noise margins of CMOS with the high current-driving capability of bipolar transistors. Specifically, CMOS, although a nearly ideal logic-circuit technology in many respects, has a limited current-driving capability. This is not a serious problem when the CMOS gate has to drive a few other CMOS gates. It becomes a serious issue, however, when relatively large capacitive loads (e.g., greater than 0.5 pF or so) are present. In such cases, one has to either resort to the use of elaborate CMOS buffer circuits or face the usually unacceptable consequence of long propagation delays. On the other hand, we know that by virtue of its much larger transconductance, the BJT is capable of large output currents. We have seen a practical illustration of that in the emitter–follower output stage of ECL.

Indeed, the high current-driving capability contributes to making ECL two to five times faster than CMOS (under equivalent conditions)—of course, at the expense of high power dissipation. In summary, then, BiCMOS seeks to combine the best of the CMOS and bipolar technologies to obtain a class of circuits that is particularly useful when output currents that are higher than possible with CMOS are needed. Furthermore, since BiCMOS technology is well suited for the implementation of high-performance analog circuits (see, e.g., Section 8.3), it makes possible the realization of both analog and digital functions on the same IC chip, making the "**system on a chip**" an attainable goal. The price paid is a more complex, and hence more expensive (than CMOS) processing technology.

15.5.1 The BiCMOS Inverter

A variety of BiCMOS inverter circuits have been proposed and are in use. All of these are based on the use of *npn* transistors to increase the output current available from a CMOS inverter. This can be most simply achieved by cascading each of the Q_N and Q_P devices of the CMOS inverter with an *npn* transistor, as shown in Fig. 15.37(a). Observe that this circuit can be thought of as utilizing the pair of complementary composite MOS-BJT devices shown in Fig. 15.37(b). These composite devices[9] retain the high input impedance of the MOS transistor while in effect multiplying its rather low g_m by the β of the BJT. It is also useful to observe that the output stage formed by Q_1 and Q_2 has what is known as the **totem-pole configuration** utilized by TTL.[10]

The circuit of Fig. 15.37(a) operates as follows: When v_I is low, both Q_N and Q_2 are off while Q_P conducts and supplies Q_1 with base current, thus turning it on. Transistor Q_1 then provides a large output current to charge the load capacitance. The result is a very fast charging of the load capacitance and correspondingly a short low-to-high propagation delay, t_{PLH}. Transistor Q_1 turns off when v_O reaches a value of about $V_{DD} - V_{BE1}$, and thus the output high level is lower than V_{DD}, a disadvantage. When v_I goes high, Q_P and Q_1 turn off, and Q_N turns on, providing its drain current into the base of Q_2. Transistor Q_2 then turns on and provides a large output current that quickly discharges the load capacitance. Here again the result is a short high-to-low propagation delay, t_{PHL}. On the negative side, Q_2 turns off when v_O reaches a value of about V_{BE2}, and thus the output low level is greater than zero, a disadvantage.

Thus, while the circuit of Fig. 15.37(a) features large output currents and short propagation delays, it has the disadvantage of reduced logic swing and, correspondingly, reduced noise margins. There is also another and perhaps more serious disadvantage, namely, the relatively long turn-off delays of Q_1 and Q_2 arising from the absence of circuit paths along which the base charge can be removed. This problem can be solved by adding a resistor between the base of each of Q_1 and Q_2 and ground, as shown in Fig. 15.37(c). Now when either Q_1 or Q_2 is turned off, its stored base charge is removed to ground through R_1 or R_2, respectively. Resistor R_2 provides an additional benefit: With v_I high, and after Q_2 cuts off, v_O continues to fall below V_{BE2}, and the output node is pulled to ground through the series path of Q_N and R_2. Thus R_2 functions as a pull-down resistor. The Q_N–R_2 path, however, is a high-impedance one with the result that pulling v_O to ground is a rather slow process. Incorporating the resistor R_1, however, is disadvantageous from a static power-dissipation standpoint: When v_I is low, a dc path exists between V_{DD} and ground through the conducting Q_P and R_1. Finally, it should

[9]It is interesting to note that these composite devices were proposed as early as 1969 (see Lin et al., 1969).

[10]Refer to the book's website for a description of the basic TTL logic-gate circuit and its totem-pole output stage.

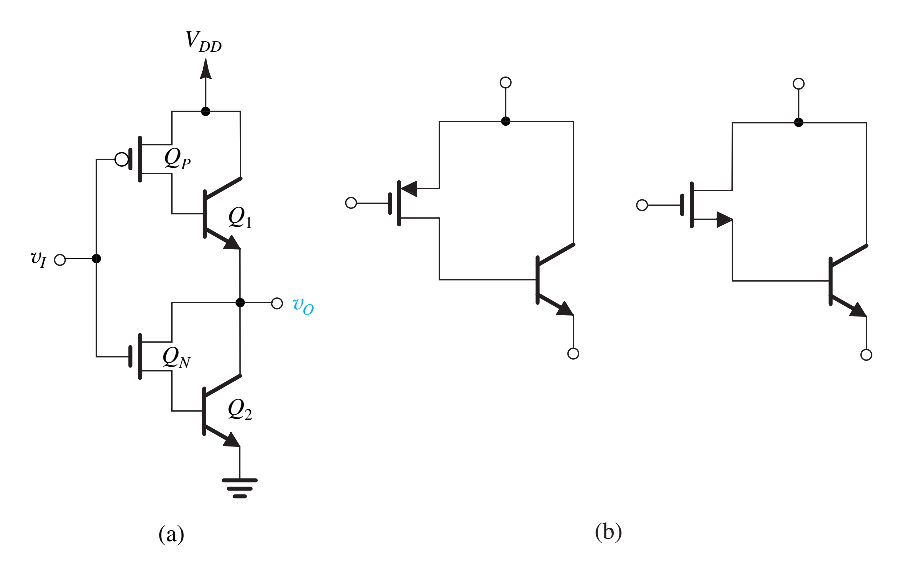


Figure 15.37 Development of the BiCMOS inverter circuit. **(a)** The basic concept is to use an additional bipolar transistor to increase the output current drive of each of Q_N and Q_P of the CMOS inverter. **(b)** The circuit in **(a)** can be thought of as utilizing these composite devices. **(c)** To reduce the turn-off times of Q_1 and Q_2, "bleeder resistors" R_1 and R_2 are added. **(d)** Implementation of the circuit in **(c)** using NMOS transistors to realize the resistors. **(e)** An improved version of the circuit in **(c)** obtained by connecting the lower end of R_1 to the output node.

be noted that R_1 and R_2 take some of the drain currents of Q_P and Q_N away from the bases of Q_1 and Q_2 and thus slightly reduce the gate output current available to charge and discharge the load capacitance.

Figure 15.37(d) shows the way in which R_1 and R_2 are usually implemented. As indicated, NMOS devices Q_{R1} and Q_{R2} are used to realize R_1 and R_2. As an added innovation, these two transistors are made to conduct only when needed. Thus, Q_{R1} will conduct only when v_I rises, at which time its drain current constitutes a reverse base current for Q_1, speeding up its turn-off. Similarly, Q_{R2} will conduct only when v_I falls and Q_P conducts, pulling the gate of Q_{R2} high. The drain current of Q_{R2} then constitutes a reverse base current for Q_2, speeding up its turn-off.

As a final circuit for the BiCMOS inverter, we show the so-called *R*-circuit in Fig. 15.37(e). This circuit differs from that in Fig. 15.37(c) in only one respect: Rather than returning R_1 to ground, we have connected R_1 to the output node of the inverter. This simple change has two benefits. First, the problem of static power dissipation is now solved. Second, R_1 now functions as a pull-up resistor, pulling the output node voltage up to V_{DD} (through the conducting Q_P) after Q_1 has turned off. Thus, the *R* circuit in Fig. 15.37(e) does in fact have output levels very close to V_{DD} and ground.

As a final remark on the BiCMOS inverter, we note that the circuit is designed so that transistors Q_1 and Q_2 are never simultaneously conducting and neither is allowed to saturate. Unfortunately, sometimes the resistance of the collector region of the BJT in conjunction with large capacitive-charging currents causes saturation to occur. Specifically, at large output currents, the voltage developed across r_C (which can be of the order of 100 Ω) can lower the voltage at the intrinsic collector terminal and cause the CBJ to become forward biased. As the reader will recall, saturation is a harmful effect for two reasons: It limits the collector current to a value less than βI_B, and it slows down the transistor turn-off.

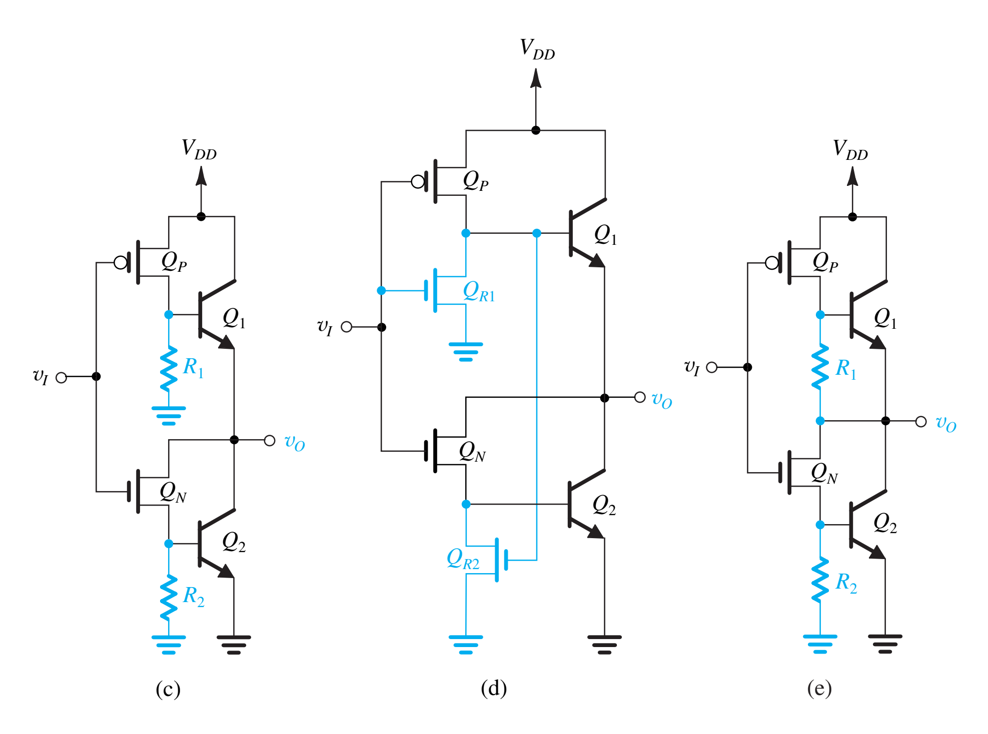


Figure 15.37 *continued*

15.5.2 Dynamic Operation

A detailed analysis of the dynamic operation of the BiCMOS inverter circuit is a rather complex undertaking. Nevertheless, an estimate of its propagation delay can be obtained by considering only the time required to charge and discharge a load capacitance C. Such an approximation is justified when C is relatively large and thus its effect on inverter dynamics is dominant: in other words, when we are able to neglect the time required to charge the parasitic capacitances present at internal circuit nodes. Fortunately, this is usually the case in practice, for if the load capacitance is not large, one would use the simpler CMOS inverter. In fact, it has been shown (Embabi, Bellaouar, and Elmasry, 1993) that the speed advantage of Bi-CMOS (over CMOS) becomes evident only when the gate is required to drive a large fan-out or a large load capacitance. For instance, at a load capacitance of 50 fF to 100 fF, BiCMOS and CMOS typically feature equal delays. However, at a load capacitance of 1 pF, t_P of a BiCMOS inverter is 0.3 ns, whereas that of an otherwise comparable CMOS inverter is about 1 ns.

Finally, in Fig. 15.38, we show simplified equivalent circuits that can be employed in obtaining rough estimates of t_{PLH} and t_{PHL} of the R-type BiCMOS inverter (see Problem 15.49).

15.5.3 BiCMOS Logic Gates

In BiCMOS, the logic is performed by the CMOS part of the gate, with the bipolar portion simply functioning as an output stage. It follows that BiCMOS logic-gate circuits can be generated following the same approach used in CMOS. As an example, we show in Fig. 15.39 a BiCMOS two-input NAND gate.

As a final remark, we note that BiCMOS technology is applied in a variety of products including microprocessors, static RAMs, and gate arrays (see Alvarez, 1993).

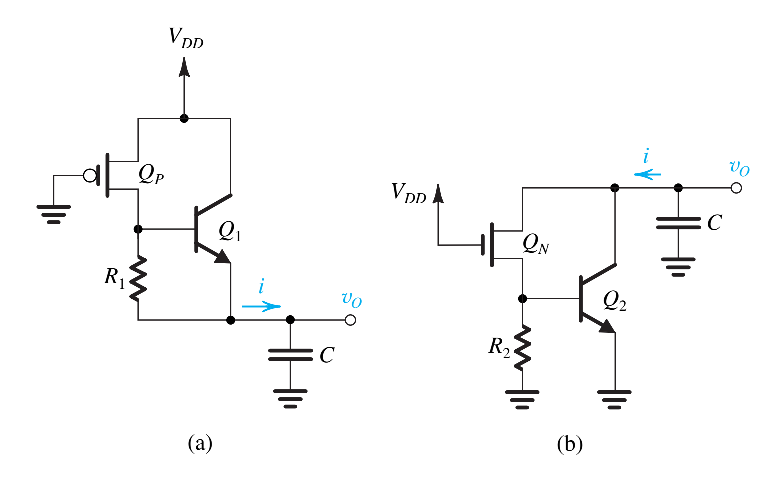


Figure 15.38 Equivalent circuits for charging and discharging a load capacitance *C*. Note that *C* includes all the capacitances present at the output node.

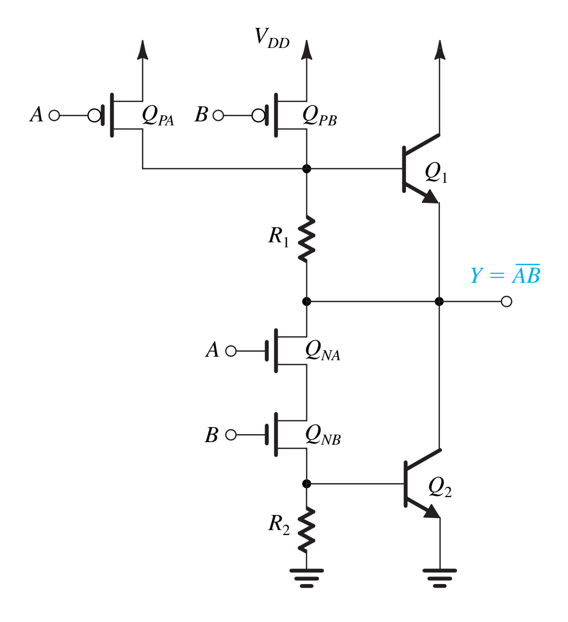


Figure 15.39 A BiCMOS two-input NAND gate.

EXERCISE

D15.17 The threshold voltage of the BiCMOS inverter of Fig. 15.37(e) is the value of v_I at which both Q_N and Q_P are conducting equal currents and operating in the saturation region. At this value of v_I, Q_2 will be on, causing the voltage at the source of Q_N to be approximately 0.7 V. It is required to design the circuit so that the threshold voltage is equal to $V_{DD}/2$. For $V_{DD} = 5$ V, $|V_t| = 0.6$ V, and assuming equal channel lengths for Q_N and Q_P and that $\mu_n \simeq 2.5\,\mu_p$, find the required ratio of widths, W_p/W_n.
Ans. 1

Section 15.4: Emitter-Coupled Logic (ECL)

D 15.34 For the ECL circuit in Fig. P15.34, the transistors exhibit V_{BE} of 0.75 V at an emitter current I and have very high β.

(a) Find V_{OH} and V_{OL}.
(b) For the input at B that is sufficiently negative for Q_B to be cut off, what voltage at A causes a current of $I/2$ to flow in Q_R?
(c) Repeat (b) for a current in Q_R of 0.99I.
(d) Repeat (c) for a current in Q_R of 0.01I.
(e) Use the results of (c) and (d) to specify V_{IL} and V_{IH}.
(f) Find NM_H and NM_L.
(g) Find the value of IR that makes the noise margins equal to the width of the transition region, $V_{IH}–V_{IL}$.

(h) Using the IR value obtained in (g), give numerical values for V_{OH}, V_{OL}, V_{IH}, V_{IL}, and V_R for this ECL gate.

***15.35** Three logic inverters are connected in a ring. Specifications for this family of gates indicate a typical propagation delay of 3 ns for high-to-low output transitions and 7 ns for low-to-high transitions. Assume that for some reason the input to one of the gates undergoes a low-to-high transition. By sketching the waveforms at the outputs of the three gates and keeping track of their relative positions, show that the circuit functions as an oscillator. What is the frequency of oscillation of this ring oscillator? In each cycle, how long is the output high? low?

***15.36** Following the idea of a ring oscillator introduced in Problem 15.35, consider an implementation using a ring of five ECL 100K inverters. Assume that the inverters have linearly rising and falling edges (and thus the waveforms are trapezoidal in shape). Let the 0 to 100% rise and fall times be equal to 1 ns. Also, let the propagation delay (for both transitions) be equal to 1 ns. Provide a labeled sketch of the five output signals, taking care that relevant phase information is provided. What is the frequency of oscillation?

D *15.37 Using the logic and circuit flexibility of ECL indicated by Figs. 15.26 and 15.36, sketch an ECL logic circuit that realizes the exclusive OR function, $Y = \bar{A}B + A\bar{B}$. Give a logic diagram (as opposed to a circuit diagram).

***15.38** For the circuit in Fig. 15.28 whose transfer characteristic is shown in Fig. 15.29, calculate the incremental voltage gain from input to the OR output at points x, m, and y of the transfer characteristic. Assume $\beta = 100$. Use the results of Exercise 15.14, and let the output at x be –1.77 V and that at y be –0.88 V. (*Hint*: Recall that x and y are defined by a 1%, 99% current split.)

15.39 For the circuit in Fig. 15.28 whose transfer characteristic is shown in Fig. 15.29, find V_{IL} and V_{IH} if x and y are defined as the points at which

(a) 90% of the current I_E is switched.
(b) 99.9% of the current I_E is switched.

15.40 For the symmetrically loaded circuit of Fig. 15.28 and for typical output signal levels ($V_{OH} = -0.88$ V and

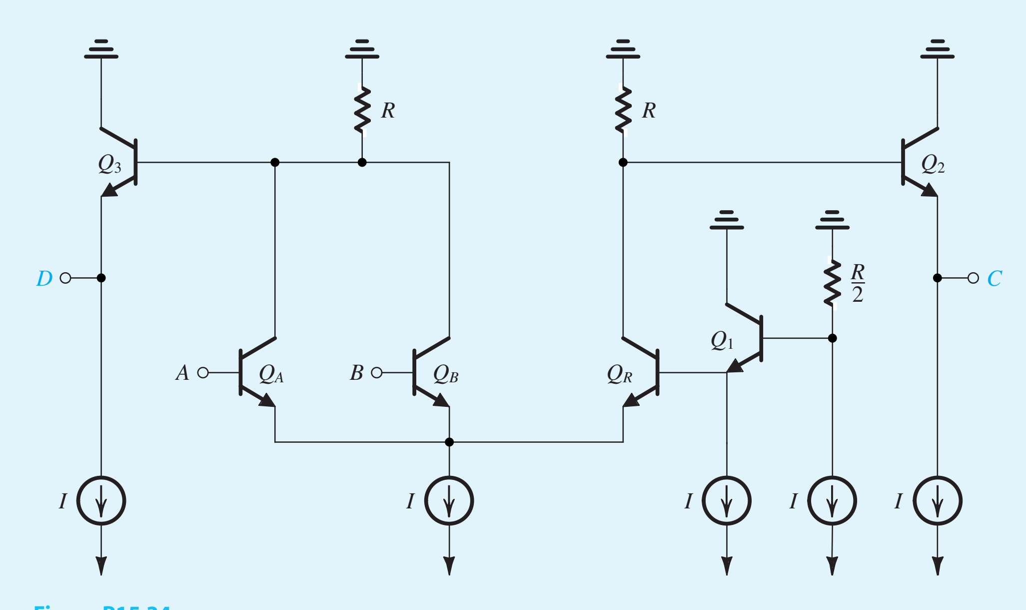


Figure P15.34

$V_{OL} = -1.77$ V), calculate the power lost in both load resistors R_T and both output followers. What then is the total power dissipation of a single ECL gate, including its symmetrical output terminations?

15.41 Considering the circuit of Fig. 15.30, what is the value of β of Q_2, for which the high noise margin (NM_H) is reduced by 50%?

***15.42** Consider an ECL gate whose inverting output is terminated in a 50-Ω resistance connected to a –2-V supply. Let the total load capacitance be denoted *C*. As the input of the gate rises, the output emitter follower cuts off and the load capacitance *C* discharges through the 50-Ω load (until the emitter follower conducts again). Find the value of *C* that will result in a discharge time of 1 ns. Assume that the two output levels are –0.88 V and –1.77 V.

15.43 For signals whose rise and fall times are 3.5 ns, what length of unterminated gate-to-gate wire interconnect can be used if a ratio of rise time to return time of 5 to 1 is required? Assume the environment of the wire to be such that the signal propagates at two-thirds the speed of light (which is 30 cm/ns).

***15.44** For the circuit in Fig. P15.44, let the levels of the inputs *A*, *B*, *C*, and *D* be 0 and +5 V. For all inputs low at 0 V, what is the voltage at *E*? If *A* and *C* are raised to +5 V, what is the voltage at *E*? Assume $|V_{BE}| = 0.7$ V and $\beta = 50$. Express *E* as a logic function of *A*, *B*, *C*, and *D*.

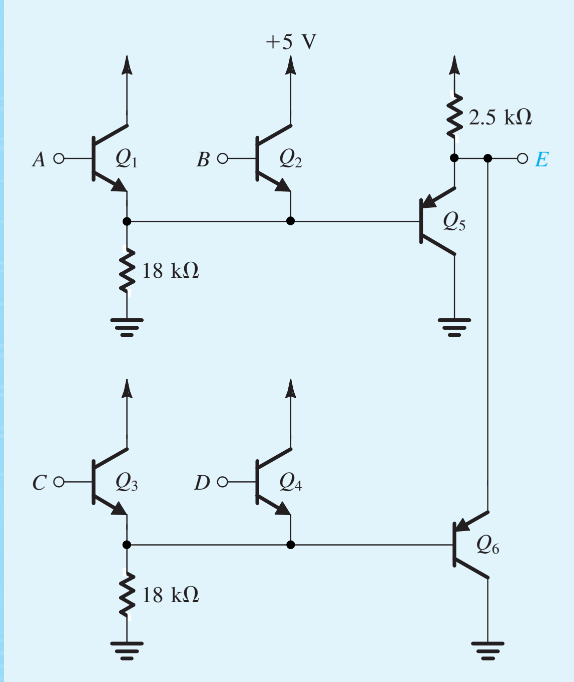


Figure P15.44

Section 15.5: BiCMOS Digital Circuits

15.45 Consider the conceptual BiCMOS circuit of Fig. 15.37(a), for the conditions that $V_{DD} = 5$ V, $|V_t| = 1$ V, $V_{BE} = 0.7$ V, $\beta = 100$, $k_n' = 2.5k_p' = 100\,\mu A/V^2$, and $(W/L)_n = 2\,\mu m/1\,\mu m$. For $v_I = v_O = V_{DD}/2$, find $(W/L)_p$ so that $I_{EQ1} = I_{EQ2}$. What is this totem-pole transient current?

15.46 Consider the conceptual BiCMOS circuit of Fig. 15.37(a) for the conditions stated in Problem 15.45. What is the threshold voltage of the inverter if both Q_N and Q_P have $W/L = 2\,\mu m/1\,\mu m$? What totem-pole current flows at v_I equal to the threshold voltage?

D *15.47 Consider the choice of values for R_1 and R_2 in the circuit of Fig. 15.37(c). An important consideration in making this choice is that the loss of base drive current will be limited. This loss becomes particularly acute when the current through Q_N and Q_P becomes small. This in turn happens near the end of the output signal swing when the associated MOS device is deeply in triode operation (say at $|v_{DS}| = |V_t|/3$). Determine values for R_1 and R_2 so that the loss in base current is limited to 50%. What is the ratio R_1/R_2? Repeat for a 20% loss in base drive.

***15.48** For the circuit of Fig. 15.37(a) with parameters as in Problem 15.45 and with $(W/L)_p = (W/L)_n$, estimate the propagation delays t_{PLH}, t_{PHL} and t_P obtained for a load capacitance of 2 pF. Assume that the internal node capacitances do not contribute much to this result. Use average values for the charging and discharging currents.

***15.49** Repeat Problem 15.48 for the circuit in Fig. 15.37(e), assuming that $R_1 = R_2 = 5$ kΩ.

D 15.50 Consider the dynamic response of the NAND gate of Fig. 15.39 with a large external capacitive load. If the worst-case response is to be identical to that of the inverter of Fig. 15.37(e), how must the *W/L* ratios of Q_{NA}, Q_{NB}, Q_N, Q_{PA}, Q_{PB}, and Q_P be related?

D 15.51 Sketch the circuit of a BiCMOS two-input NOR gate. If, when loaded with a large capacitance, the gate is to have worst-case delays equal to the corresponding values of the inverter of Fig. 15.37(e), find *W/L* of each transistor in terms of $(W/L)_n$ and $(W/L)_p$.

18.9 Precision Rectifier Circuits

Rectifier circuits were studied in Chapter 4, where the emphasis was on their application in power-supply design. In such applications, the voltages being rectified are usually much greater than the diode voltage drop, rendering the exact value of the diode drop unimportant to the proper operation of the rectifier. Other applications exist, however, where this is not the case. For instance, in instrumentation applications, the signal to be rectified can be of a very small amplitude, say 0.1 V, making it impossible to employ the conventional rectifier circuits. Also, in instrumentation applications, the need arises for rectifier circuits with very precise transfer characteristics.

Here we study circuits that combine diodes and op amps to implement a variety of rectifier circuits with precise characteristics. Precision rectifiers, which can be considered a special class of wave-shaping circuits, find application in the design of instrumentation systems. An introduction to precision rectifiers was presented in Section 4.5.5. This material, however, is repeated here for the reader's convenience.

18.9.1 Precision Half-Wave Rectifier: The "Superdiode"

Figure 18.35(a) shows a precision half-wave-rectifier circuit consisting of a diode placed in the negative-feedback path of an op amp, with R being the rectifier load resistance. The circuit works as follows: If v_I goes positive, the output voltage v_A of the op amp will go positive and the diode will conduct, thus establishing a closed feedback path between the op amp's output terminal and the negative input terminal. This negative-feedback path will cause a virtual short circuit to appear between the two input terminals of the op amp. Thus the voltage at the negative input terminal, which is also the output voltage v_O, will equal (to within a few millivolts) that at the positive input terminal, which is the input voltage v_I,

$$v_O = v_I \quad v_I \geq 0$$

Note that the offset voltage ($\simeq 0.5$ V) exhibited in the simple half-wave-rectifier circuit is no longer present. For the op-amp circuit to start operation, v_I has to exceed only a negligibly small voltage equal to the diode drop divided by the op amp's open-loop gain. In other words, the straight-line transfer characteristic v_O–v_I almost passes through the origin. This makes this circuit suitable for applications involving very small signals.

Consider now the case when v_I goes negative. The op amp's output voltage v_A will tend to follow and go negative. This will reverse-bias the diode, and no current will flow through resistance R, so that v_O remains equal to 0 V. Thus for $v_I < 0$, $v_O = 0$. Since in this case the diode is off, the op amp will be operating in an open-loop fashion and its output will be at the negative saturation level.

The transfer characteristic of this circuit will be that shown in Fig. 18.35(b), which is almost identical to the ideal characteristic of a half-wave rectifier. The nonideal diode characteristics have been almost completely masked by placing the diode in the negative-feedback path of an op amp. This is another dramatic application of negative feedback. The combination of diode and op amp, shown in the dashed box in Fig. 18.35(a), is appropriately referred to as a "superdiode."

As usual, though, not all is well. The circuit of Fig. 18.35 has some disadvantages: When v_I goes negative and $v_O = 0$, the entire magnitude of v_I appears between the two input terminals of the op amp. If this magnitude is greater than a few volts, the op amp may be damaged unless it is equipped with what is called "overvoltage protection" (a feature that most modern IC

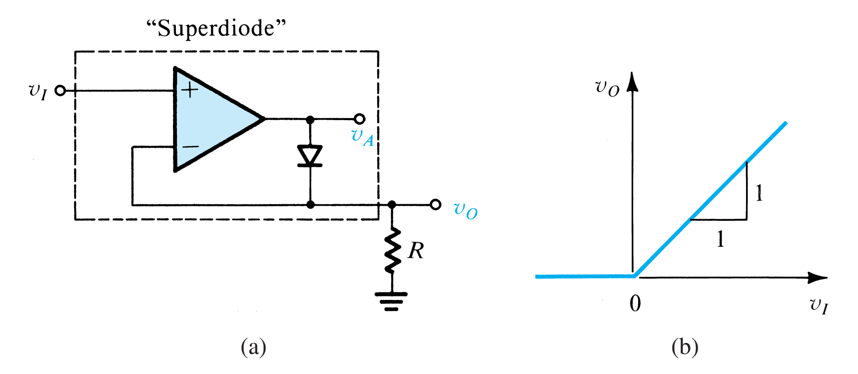


Figure 18.35 **(a)** The "superdiode" precision half-wave rectifier; **(b)** its almost ideal transfer characteristic. Note that when $v_I > 0$ and the diode conducts, the op amp supplies the load current, and the source is conveniently buffered, an added advantage.

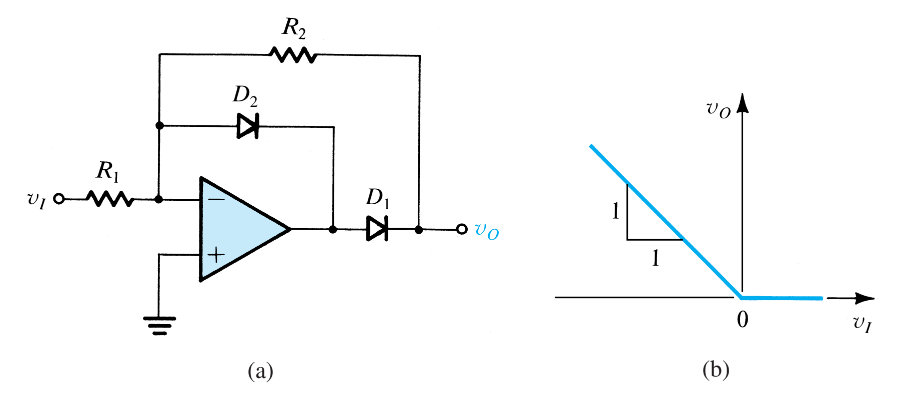


Figure 18.36 **(a)** An improved version of the precision half-wave rectifier: Diode D_2 is included to keep the feedback loop closed around the op amp during the off times of the rectifier diode D_1, thus preventing the op amp from saturating. **(b)** The transfer characteristic for $R_2 = R_1$.

op amps have). Another disadvantage is that when v_I is negative, the op amp will be saturated. Although not harmful to the op amp, saturation should usually be avoided, since getting the op amp out of the saturation region and back into its linear region of operation requires some time. This time delay will obviously slow down circuit operation and limit the frequency of operation of the superdiode half-wave-rectifier circuit.

18.9.2 An Alternative Circuit

An alternative precision rectifier circuit that does not suffer from the disadvantages mentioned above is shown in Fig. 18.36. The circuit operates in the following manner: For positive v_I, diode D_2 conducts and closes the negative-feedback loop around the op amp. A virtual ground therefore will appear at the inverting input terminal, and the op amp's output will be *clamped* at one diode drop below ground. This negative voltage will keep diode D_1 off, and no current will flow in the feedback resistance R_2. It follows that the rectifier output voltage will be zero.

As v_I goes negative, the voltage at the inverting input terminal will tend to go negative, causing the voltage at the op amp's output terminal to go positive. This will cause D_2 to be

reverse-biased and hence to be cut off. Diode D_1, however, will conduct through R_2, thus establishing a negative-feedback path around the op amp and forcing a virtual ground to appear at the inverting input terminal. The current through the feedback resistance R_2 will be equal to the current through the input resistance R_1. Thus for $R_1 = R_2$ the output voltage v_O will be

$$v_O = -v_I \quad v_I \leq 0$$

The transfer characteristic of the circuit is shown in Fig. 18.36(b). Note that unlike the situation for the circuit shown in Fig. 18.35, here the slope of the characteristic can be set to any desired value, including unity, by selecting appropriate values for R_1 and R_2.

As mentioned before, the major advantage of the improved half-wave-rectifier circuit is that the feedback loop around the op amp remains closed at all times. Hence the op amp remains in its linear operating region, avoiding the possibility of saturation and the associated time delay required to "get out" of saturation. Diode D_2 "catches" the op-amp output voltage as it goes negative and clamps it to one diode drop below ground; hence D_2 is called a "catching diode."

18.9.3 An Application: Measuring AC Voltages

As one of the many possible applications of the precision rectifier circuits discussed in this section, consider the basic ac voltmeter circuit shown in Fig. 18.37. The circuit consists of a half-wave rectifier—formed by op amp A_1, diodes D_1 and D_2, and resistors R_1 and R_2—and a first-order low-pass filter—formed by op amp A_2, resistors R_3 and R_4, and capacitor C. For an input sinusoid having a peak amplitude V_p the output v_1 of the rectifier will consist of a half sine wave having a peak amplitude of V_pR_2/R_1. It can be shown using Fourier series analysis that the waveform of v_1 has an average value of $(V_p/\pi)(R_2/R_1)$ in addition to

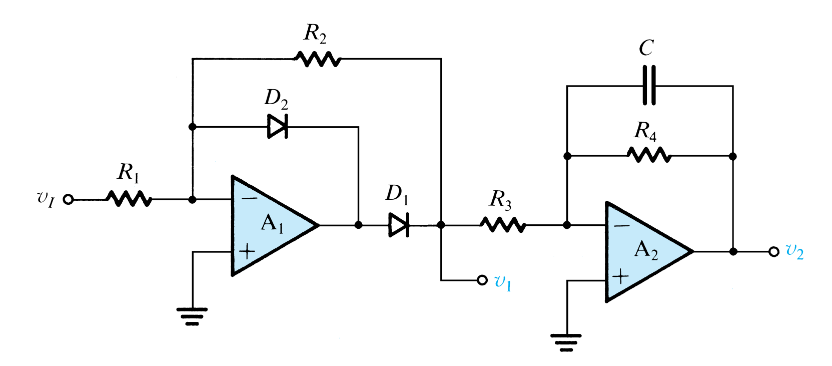


Figure 18.37 A simple ac voltmeter consisting of a precision half-wave rectifier followed by a first-order low-pass filter.

EXERCISES

18.27 Consider the operational rectifier or superdiode circuit of Fig. 18.35(a), with $R = 1\ k\Omega$. For $v_I = 10$ mV, 1 V, and –1 V, what are the voltages that result at the rectifier output and at the output of the op amp?

Assume that the op amp is ideal and that its output saturates at ±12 V. The diode has a 0.7-V drop at 1-mA current, and the voltage drop changes by 0.1 V per decade of current change.
Ans. 10 mV, 0.51 V; 1 V, 1.7 V; 0 V, –12 V

18.28 If the diode in the circuit of Fig. 18.35(a) is reversed, what is the transfer characteristic v_O as a function of v_I?
Ans. $v_O = 0$ for $v_I \geq 0$; $v_O = v_I$ for $v_I \leq 0$

18.29 Consider the circuit in Fig. 18.36(a) with $R_1 = 1$ kΩ and $R_2 = 10$ kΩ. Find v_O and the voltage at the amplifier output for $v_I = +1$ V, –10 mV, and –1 V. Assume the op amp to be ideal with saturation voltages of ±12 V. The diodes have 0.7-V voltage drops at 1 mA, and the voltage drop changes by 0.1 V per decade of current change.
Ans. 0 V, −0.vm7 V; 0.1 V, 0.6 V; 10 V, 10.7 V

18.30 If the diodes in the circuit of Fig. 18.36(a) are reversed, what is the transfer characteristic v_O as a function of v_I?
Ans. $v_O = -(R_2/R_1)v_I$ for $v_I \geq 0$; $v_O = 0$ for $v_I \leq 0$

18.31 Find the transfer characteristic for the circuit in Fig. E18.31.

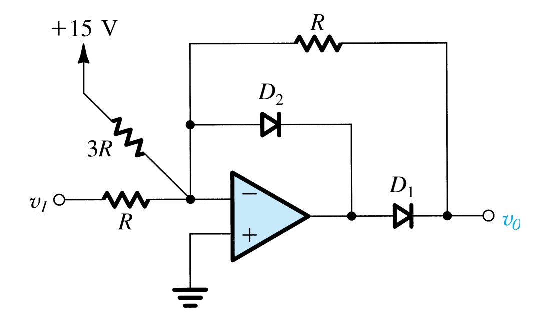


Figure E18.31

Ans. $v_O = 0$ for $v_I \geq -5$V; $v_O = -v_I - 5$ for $v_I \leq -5$ V

harmonics of the frequency ω of the input signal. To reduce the amplitudes of all these harmonics to negligible levels, the corner frequency of the low-pass filter should be chosen to be much smaller than the lowest expected frequency ω_{min} of the input sine wave. This leads to

$$\frac{1}{CR_4} \ll \omega_{min}$$

Then the output voltage v_2 will be mostly dc, with a value

$$V_2 = -\frac{V_p}{\pi}\frac{R_2}{R_1}\frac{R_4}{R_3}$$

where R_4/R_3 is the dc gain of the low-pass filter. Note that this voltmeter essentially measures the average value of the negative parts of the input signal but can be calibrated to provide rms readings for input sinusoids.

18.9.4 Precision Full-Wave Rectifier

We now derive a circuit for a precision full-wave rectifier. From Chapter 4 we know that full-wave rectification is achieved by inverting the negative halves of the input-signal waveform and applying the resulting signal to another diode rectifier. The outputs of the two rectifiers are then joined to a common load. Such an arrangement is depicted in Fig. 18.38, which also shows the waveforms at various nodes. Now replacing diode D_A with a superdiode, and replacing diode D_B and the inverting amplifier with the inverting precision half-wave rectifier of Fig. 18.36 but without the catching diode, we obtain the precision full-wave-rectifier circuit of Fig. 18.39(a).

To see how the circuit of Fig. 18.39(a) operates, consider first the case of positive input at A. The output of A_2 will go positive, turning D_2 on, which will conduct through R_L and thus close the feedback loop around A_2. A virtual short circuit will thus be established between the two input terminals of A_2, and the voltage at the negative-input terminal, which is the output voltage of the circuit, will become equal to the input. Thus no current will flow through R_1 and R_2, and the voltage at the inverting input of A_1 will be equal to the input and hence positive. Therefore the output terminal (F) of A_1 will go negative until A_1 saturates. This causes D_1 to be turned off.

Next consider what happens when A goes negative. The tendency for a negative voltage at the negative input of A_1 causes F to rise, making D_1 conduct to supply R_L and allowing the feedback loop around A_1 to be closed. Thus a virtual ground appears at the negative input of A_1, and the two equal resistances R_1 and R_2 force the voltage at C, which is the output

EXERCISES

18.32 In the full-wave rectifier circuit of Fig. 18.39(a), let $R_1 = R_2 = R_L = 10$ kΩ and assume the op amps to be ideal except for output saturation at ±12 V. When conducting a current of 1 mA, each diode exhibits a voltage drop of 0.7 V, and this voltage changes by 0.1 V per decade of current change. Find v_O, v_E, and v_F corresponding to $v_I = +0.1$ V, +1 V, +10 V, –0.1 V, and –10 V.
Ans. + 0.1 V, + 0.6 V, –12 V; +1 V, +1.6 V, –12 V; +10 V, +10.7 V, –12 V; + 0.1 V, –12 V, + 0.63 V; +1 V, –12 V, +1.63 V; +10 V, –12 V, +10.73 V

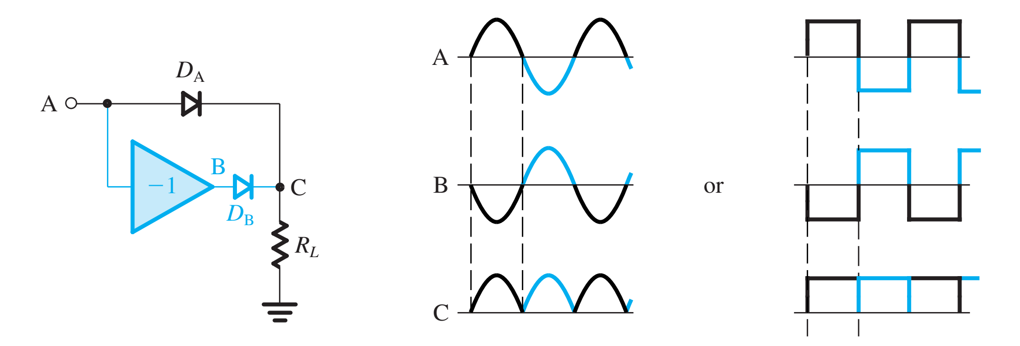


Figure 18.38 Principle of full-wave rectification.

D18.33 The block diagram shown in Fig. E18.33(a) gives another possible arrangement for implementing the absolute-value or full-wave-rectifier operation depicted symbolically in Fig. E18.33(b). The block diagram consists of two boxes: a half-wave rectifier, which can be implemented by the circuit in Fig. 18.36(a) after reversing both diodes, and a weighted inverting summer. Convince yourself that this block diagram does in fact realize the absolute-value operation. Then draw a complete circuit diagram, giving reasonable values for all resistors.

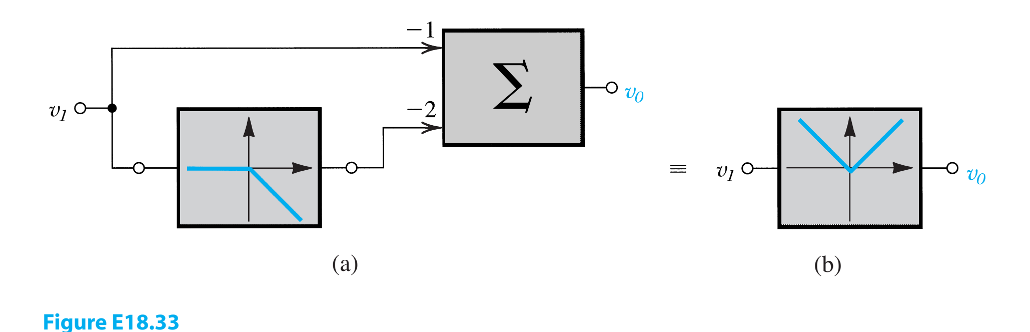


Figure E18.33

voltage, to be equal to the negative of the input voltage at A and thus positive. The combination of positive voltage at C and negative voltage at A causes the output of A_2 to saturate in the negative direction, thus keeping D_2 off.

The overall result is perfect full-wave rectification, as represented by the transfer characteristic in Fig. 18.39(b). This precision is, of course, a result of placing the diodes in op-amp feedback loops, thus masking their nonidealities. This circuit is one of many possible precision full-wave-rectifier or **absolute-value circuits**. Another related implementation of this function is examined in Exercise 18.33.

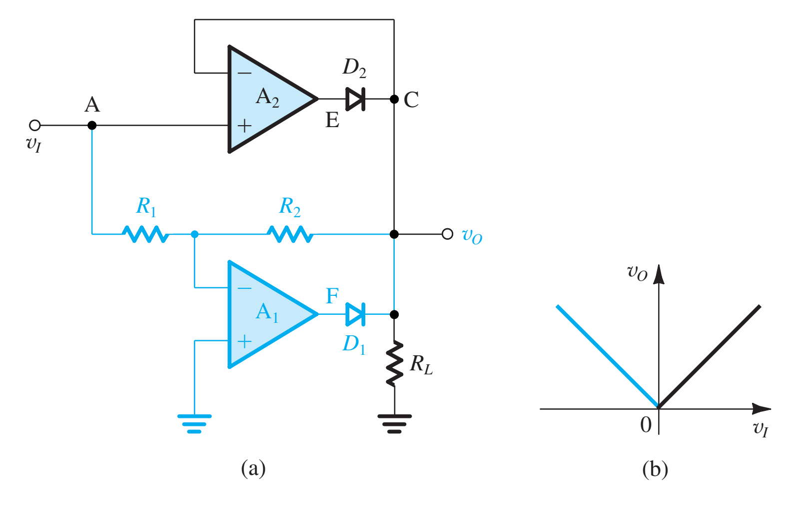


Figure 18.39 **(a)** Precision full-wave rectifier based on the conceptual circuit of Fig. 18.38. **(b)** Transfer characteristic of the circuit in **(a)**.

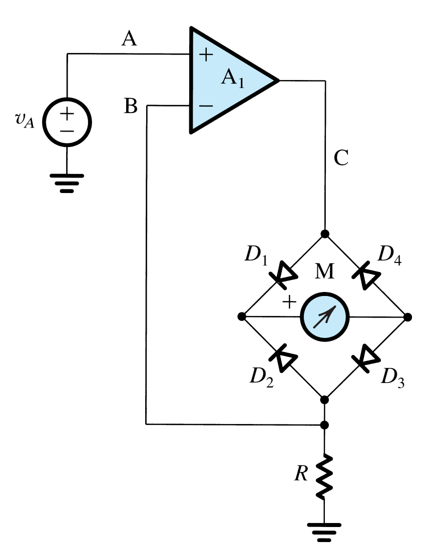


Figure 18.40 Use of the diode bridge in the design of an ac voltmeter.

18.9.5 A Precision Bridge Rectifier for Instrumentation Applications

The bridge rectifier circuit studied in Section 4.5.3 can be combined with an op amp to provide useful precision circuits. One such arrangement is shown in Fig. 18.40. This circuit causes a current equal to $|v_A|/R$ to flow through the moving-coil meter M. Thus the meter provides a reading that is proportional to the average of the absolute value of the input voltage v_A. All the nonidealities of the meter and of the diodes are masked by placing the bridge circuit in the negative-feedback loop of the op amp. Observe that when v_A is positive, current flows from the op-amp output through D_1, M, D_3, and R. When v_A is negative, current flows into the op-amp output through R, D_2, M, and D_4. Thus the feedback loop remains closed for both polarities of v_A. The resulting virtual short circuit at the input terminals of the op amp causes a replica of v_A to appear across R. The circuit of Fig. 18.40 provides a relatively accurate high-input-impedance ac voltmeter using an inexpensive moving-coil meter.

EXERCISE

D18.34 In the circuit of Fig. 18.40, find the value of R that would cause the meter to provide a full-scale reading when the input voltage is a sine wave of 5 V rms. Let meter M have a 1-mA, 50-Ω movement (i.e., its resistance is 50 Ω, and it provides full-scale deflection when the average current through it is 1 mA). What are the approximate maximum and minimum voltages at the op amp's output? Assume that the diodes have constant 0.7-V drops when conducting.
Ans. 4.5 kΩ; +8.55 V; −8.55 V

18.9.6 Precision Peak Rectifiers

Including the diode of the peak rectifier studied in Section 4.5.4 inside the negative-feedback loop of an op amp, as shown in Fig. 18.41, results in a precision peak rectifier. The diode–op-amp combination will be recognized as the superdiode of Fig. 18.35(a). Operation of the circuit in Fig. 18.41 is quite straightforward. For v_I greater than the output voltage, the op amp will drive the diode on, thus closing the negative-feedback path and causing the op amp to act as a follower. The output voltage will therefore follow that of the input, with the op amp supplying the capacitor-charging current. This process continues until the input reaches its peak value. Beyond the positive peak, the op amp will see a negative voltage between its input terminals. Thus its output will go negative to the saturation level and the diode will turn off. Except for possible discharge through the load resistance, the capacitor will retain a voltage equal to the positive peak of the input. Inclusion of a load resistance is essential if the circuit is required to detect reductions in the magnitude of the positive peak.

18.9.7 A Buffered Precision Peak Detector

When the peak detector is required to hold the value of the peak for a long time, the capacitor should be buffered, as shown in the circuit of Fig. 18.42. Here op amp A_2, which should have high input impedance and low input bias current, is connected as a voltage follower. The remainder of the circuit is quite similar to the half-wave-rectifier circuit of Fig. 18.42. While diode D_1 is the essential diode for the peak-rectification operation, diode D_2 acts as a catching diode to prevent negative saturation, and the associated delays, of op amp A_1. During the holding state, follower A_2 supplies D_2 with a small current through R. The output of op amp A_1 will then be clamped at one diode drop below the input voltage. Now if the input v_I

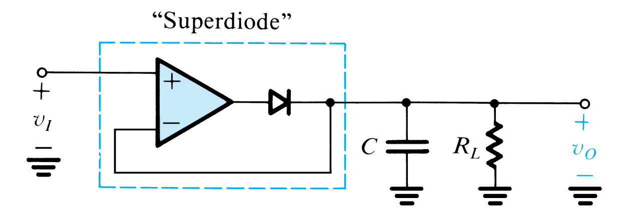


Figure 18.41 A precision peak rectifier obtained by placing the diode in the feedback loop of an op amp.

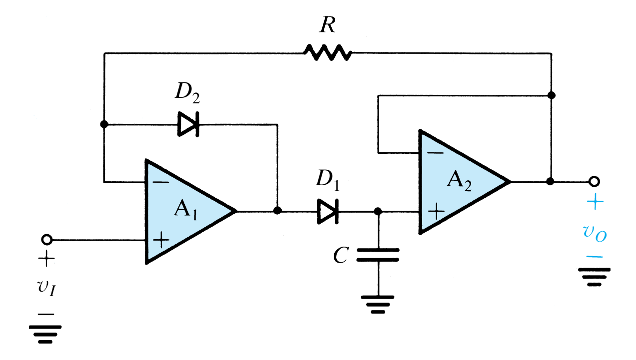


Figure 18.42 A buffered precision peak rectifier.

Figure 18.43 A precision clamping circuit.

increases above the value stored on C, which is equal to the output voltage v_O, op amp A_1 sees a net positive input that drives its output toward the positive saturation level, turning off diode D_2. Diode D_1 is then turned on and capacitor C is charged to the new positive peak of the input, after which time the circuit returns to the holding state. Finally, note that this circuit has a low-impedance output.

18.9.8 A Precision Clamping Circuit

By replacing the diode in the clamping circuit studied in Section 4.6.2 with a "superdiode," the precision clamp of Fig. 18.43 is obtained. Operation of this circuit should be self-explanatory.

PROBLEMS

Section 18.9: Precision Rectifier Circuits

18.54 Two superdiode circuits connected to a common-load resistor and having the same input signal have their diodes reversed, one with cathode to the load, the other with anode to the load. For a sine-wave input of 10 V peak to peak, what is the output waveform? Note that each half-cycle of the load current is provided by a separate amplifier, and that while one amplifier supplies the load current, the other amplifier idles. This idea, called class B operation (see Chapter 12), is important in the implementation of power amplifiers.

D 18.55 The superdiode circuit of Fig. 18.35(a) can be made to have gain by connecting a resistor R_2 in place of the short circuit between the cathode of the diode and the negative-input terminal of the op amp, and a resistor R_1 between the negative-input terminal and ground. Design the circuit for a gain of 2. For a 10-V peak-to-peak input sine wave, what is the average output voltage resulting?

D 18.56 Provide a design of the inverting precision rectifier shown in Fig. 18.36(a) in which the gain is –2 for negative inputs and zero otherwise, and the input resistance is 100 kΩ. What values of R_1 and R_2 do you choose?

D *18.57 Provide a design for a voltmeter circuit similar to the one in Fig. 18.37, which is intended to function at frequencies of 10 Hz and above. It should be calibrated for sine-wave input signals to provide an output of +10 V for an input of 1 V rms. The input resistance should be as high as possible. To extend the bandwidth of operation, keep the gain in the ac part of the circuit reasonably small. As well, the design should result in reduction of the size of the capacitor C required. The largest value of resistor available is 1 MΩ.

18.58 Plot the transfer characteristic of the circuit in Fig. P18.58.

18.59 Plot the transfer characteristics v_{O1}–v_I and v_{O2}–v_I of the circuit in Fig. P18.59.

18.60 Sketch the transfer characteristics of the circuit in Fig. P18.60.

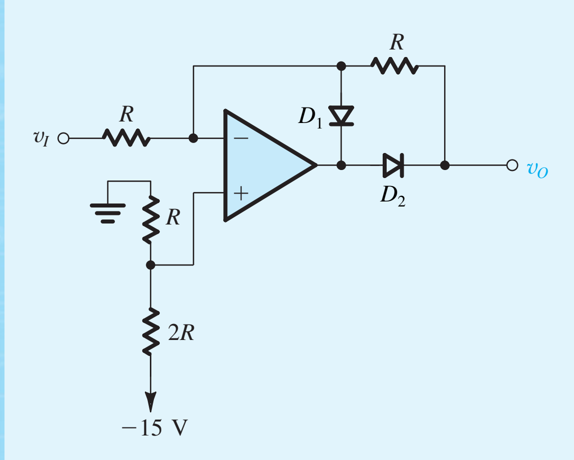


Figure P18.58

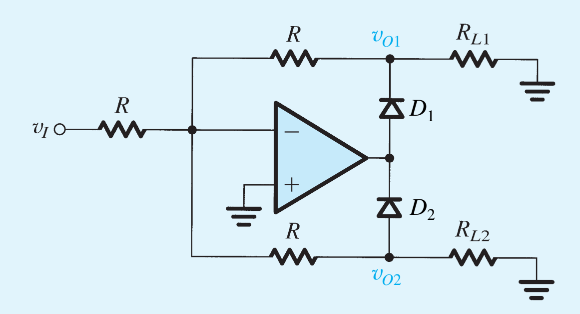


Figure P18.59

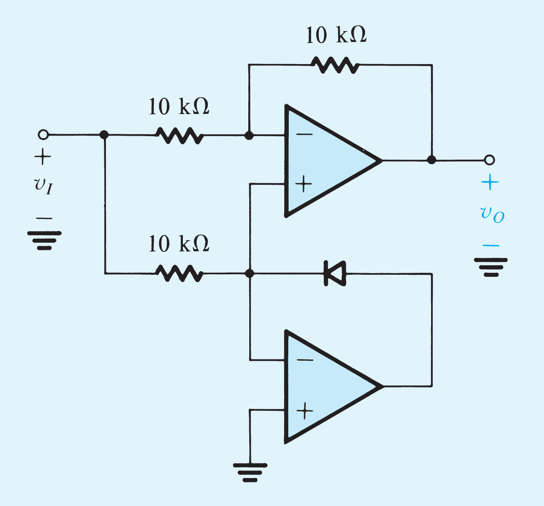


Figure P18.60

D 18.61 A circuit related to that in Fig. 18.40 is to be used to provide a current proportional to $v_A (v_A \geq 0)$ to a light-emitting diode (LED). The value of the current is to be independent of the diode's nonlinearities and variability. Indicate how this may be done easily.

***18.62** In the precision rectifier of Fig. 18.40, the resistor R is replaced by a capacitor C. What happens? For equivalent performance with a sine-wave input of 60-Hz frequency with $R = 1\ \text{k}\Omega$, what value of C should be used? What is the response of the modified circuit at 120 Hz? At 180 Hz? If the amplitude of v_A is kept fixed, what new function does this circuit perform? Now consider the effect of a waveform change on both circuits (the one with R and the one with C). For a triangular-wave input of 60-Hz frequency that produces an average meter current of 1 mA in the circuit with R, what does the average meter current become when R is replaced with the C whose value was just calculated?

***18.63** A positive-peak rectifier utilizing a fast op amp and a junction diode in a superdiode configuration, and a 10-µF capacitor initially uncharged, is driven by a series of 10-V pulses of 10-µs duration. If the maximum output current that the op amp can supply is 10 mA, what is the voltage on the capacitor following one pulse? Two pulses? Ten pulses? How many pulses are required to reach 0.5 V? 1.0 V? 2.0 V?

D 18.64 Consider the buffered precision peak rectifier shown in Fig. 18.42 when connected to a triangular input of 1-V peak-to-peak amplitude and 1000-Hz frequency. It utilizes an op amp whose bias current (directed into A_2) is 10 nA and diodes whose reverse leakage current is 1 nA. What is the smallest capacitor that can be used to guarantee an output ripple less than 1%?